

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 16:12:04 ; Search time 18 seconds
(without alignments)
546.737 Million cell updates/sec

Title: US-09-896-580B-12

Perfect score: 991

Sequence: 1 MLTNKDIIRDGHPTLRQAA.....KDHLPQHTDAVEVHQHHH 189

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	943	95.2	183	1 DEF_STAAM	Q8x78 staphylococ
2	938	94.7	183	1 DEF_STAAM	Q8f414 staphylococ
3	795	80.2	183	1 DEF_STAAP	Q8cpn4 staphylococ
4	575.5	58.1	184	1 DEF2_BACSU	Q45495 bacillus su
5	550.5	55.5	184	1 DEF2_BACCR	Q819k2 bacillus ce
6	548.5	55.3	183	1 DEF_OCEIH	Q8er96 oceanobacil
7	548.5	55.3	184	1 DEF2_BACAA	Q81mc9 bacillus an
8	540.5	54.5	184	1 DEF2_BACST	Q31410 bacillus st
9	536.5	54.1	183	1 DEF_LISIN	Q92cx8 listeria in
10	533.5	53.8	183	1 DEF_LISMO	Q8y866 listeria mo
11	506.5	51.1	182	1 DEF_BACHD	Q9k919 bacillus ha
12	488	49.2	186	1 DEF_LACPL	Q88vb2 lactobacill
13	478.5	48.3	187	1 DEF_ENTFA	Q82j10 enterococcu
14	466	47.0	204	1 DEF_STRPY	Q99xv7 streptococc
15	465	46.9	204	1 DEF_STRP8	Q8nbz7 streptococc
16	453	45.7	204	1 DEF_STRA3	Q8e378 streptococc
17	453	45.7	204	1 DEF_STRA5	Q8dx16 streptococc
18	451.5	45.6	203	1 DEF_STRR6	Q8dp79 streptococc
19	450.5	45.5	203	1 DEF_STRPN	Q8df20 streptococc
20	435	43.9	204	1 DEF_STRMU	Q8dwc2 streptococc
21	426	43.0	196	1 DEF_LACLA	Q48661 lactococcus
22	273.5	27.6	198	1 DEF_URPFA	Q9pq25 ureaplasma
23	270	27.2	186	1 DEF_MYCP8	Q8evj8 mycoplasma
24	268.5	27.1	213	1 DEF_MYCPA	Q7nak8 mycoplasma
25	263	26.5	198	1 DEF_MYCPU	Q75527 mycoplasma
26	228.5	23.1	216	1 DEF_PASMU	P57348 pasteurella
27	227	22.9	170	1 DEF_RICCN	Q92j17 rickettsia
28	222.5	22.5	202	1 DEF2_RICCN	Q7mcq2 vibrio vuln
29	221.5	22.4	216	1 DEF_MYCCE	Q7352 mycoplasma
30	219.5	22.1	168	1 DEF_VIBVY	Q7mcq2 vibrio vuln
31	218.5	22.0	168	1 DEF2_VIBCH	Q8nl6 vibrio chol
32	218.5	22.0	168	1 DEF2_VIBVU	Q8dps5 vibrio vuln
33	217.5	21.9	179	1 DEF2_ANASP	Q8yvn1 anabaena sp

ALIGNMENTS

RESULT 1

ID	DEF_STAAM	STANDARD;	PRT;	183 AA.
AC	Q8X78;			
DT	15-MAR-2004	(Rel. 43, Created)		
DT	15-MAR-2004	(Rel. 43, Last sequence update)		
DT	15-MAR-2004	(Rel. 43, Last annotation update)		
DE	Peptide deformylase (EC 3.5.1.88) (PDF) (Polypeptide deformylase).			
GN	DEF OR MW0974.			
OS	Staphylococcus aureus (strain MW2).			
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.			
OX	NCBI_TaxID=196620;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=22040717; PubMed=12044378;			
RA	Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,			
RA	Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Chui L.,			
RA	Yamamoto K., Hiramatsu K.;			
RT	"Genome and virulence determinants of high virulence community-			
RT	acquired MRSA.";			
RL	Lancet 359:1819-1827(2002).			
CC	-1- FUNCTION: Removes the formyl group from the N-terminal Met of			
CC	newly synthesized proteins. Requires at least a dipeptide for an			
CC	efficient rate of reaction. N-terminal L-methionine is a			
CC	prerequisite for activity but the enzyme has broad specificity at			
CC	other positions (By similarity).			
CC	-1- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +			
CC	methionyl peptide.			
CC	-1- COFACTOR: Binds 1 iron(II) ion (By similarity).			
CC	-1- SIMILARITY: Belongs to the polypeptide deformylase family.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	EMBL; AF004825; BAB94839.1; --			
DR	HMAP; MF 00163; -- 1.			
DR	InterPro; IPR000181; Fmet deformylase.			
DR	Pfam; PF01327; Pep_Deformylase; 1.			
DR	PRINTS; PR01576; PDEFORMYLASE.			
DR	ProDom; PD003844; Pep deformylase; 1.			
DR	TIGRFAMs; TIGR00079; pep deformyl; 1.			
KW	Protein biosynthesis; Hydrolase; Iron; Complete proteome.			
FT	ACT SITE 155 155			
FT	METAL 111 111 IRON (BY SIMILARITY).			
FT	METAL 154 154 IRON (BY SIMILARITY).			
FT	METAL 158 158 IRON (BY SIMILARITY).			
SQ	SEQUENCE 183 AA; 20559 MW; 32A64066AE5CAB0E CRC64;			

Query Match 95.2%; Score 943; DB 1; Length 183;

Best Local Similarity 99.5%; Pred. No. 8.9e-72;

Matches 182; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MLTMKDIIRDGHPTLRQAAELPLTKKEKTLIAMREFLVNSQDEETAKRYGLRSGVG 60
DB 1 MLTMKDIIRDGHPTLRQAAELPLTKKEKTLIAMREFLVNSQDEETAKRYGLRSGVG 60
QY 61 LAAPQINISKRMIAVLIPDDGSGKSYDYMLVNPVKIVSHSVQAYLPTGEGCLSVDDNVAG 120
DB 61 LAAPQINISKRMIAVLIPDDGSGKSYDYMLVNPVKIVSHSVQAYLPTGEGCLSVDDNVAG 120
QY 121 LVHRNKITIKAKDIEGNDIQLRLKGYPAIVFQHEIDLNGVMFYDHDKHPLOPHPTDA 180
DB 121 LVHRNRITIKAKDIEGNDIQLRLKGYPAIVFQHEIDLNGVMFYDHDKHPLOPHPTDA 180
QY 181 VEV 183
DB 181 VEV 183

RESULT 2
DEF STAAM STANDARD; PRT; 183 AA.
AC Q9F4L4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Peptide deformylase (EC 3.5.1.88) (PDF) (Polypeptide deformylase).
GN DEF OR DEF1 OR PDF1 OR SAVI091 OR SA0942.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879, 1280;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES-S. aureus; STRAIN=WCUH29;
RA Lonetto M.A., Sylvester D.R., Warren R.L.;
RT "Staphylococcus aureus deformylase 1 encoding DNA.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MU50 / ATCC 700699, and N315;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Lui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hoshiyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yanashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
aureus.";
RL Lancet 357:1225-1240(2001).
CC -!- FUNCTION: Removes the formyl group from the N-terminal Met of
CC newly synthesized proteins. Requires at least a dipeptide for an
CC efficient rate of reaction. N-terminal L-methionine is a
CC prerequisite for activity but the enzyme has broad specificity at
CC other positions (By similarity).
CC -!- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +
CC methionyl peptide.
CC -!- COFACTOR: Binds 1 iron(II) ion (By similarity).
CC -!- SIMILARITY: Belongs to the polypeptide deformylase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AY007227; AAG0249.1; -
CC EMBL; AP003361; BAB57253.1; -
CC EMBL; AP003132; BAB42188.1; -

```

```

DR PIR; A89879; A89879.
DR PDB; 1LOW; 24-JUL-02.
DR SWISS-2DPAGE; Q9F4L4; STAA.
DR HAMAP; MF 00163; -; 1.
DR InterPro; IPR000181; Pep deformylase.
DR Pfam; PF01327; Pep deformylase; 1.
DR PRINTS; PR01576; PDEFORMLASE.
DR PRODOM; PD003844; Pep deformylase; 1.
DR TIGRFAMs; TIGR00079; pep deformyl; 1.
KW Protein biosynthesis; Hydrolase; Iron; Complete proteome;
KW 3D-structure.
FT ACT SITE 155 155 BY SIMILARITY.
FT METAL 111 111 IRON (BY SIMILARITY).
FT METAL 154 154 IRON (BY SIMILARITY).
FT METAL 158 158 IRON (BY SIMILARITY).
SQ SEQUENCE 183 AA; 20558 MW; 32A64066A6FEAB0E CRC64;

Query Match 94.7%; Score 938; DB 1; Length 183;
Best Local Similarity 98.9%; Pred. No. 2.3e-71;
Matches 181; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLTMKDIIRDGHPTLRQAAELPLTKKEKTLIAMREFLVNSQDEETAKRYGLRSGVG 60
DB 1 MLTMKDIIRDGHPTLRQAAELPLTKKEKTLIAMREFLVNSQDEETAKRYGLRSGVG 60
QY 61 LAAPQINISKRMIAVLIPDDGSGKSYDYMLVNPVKIVSHSVQAYLPTGEGCLSVDDNVAG 120
DB 61 LAAPQINISKRMIAVLIPDDGSGKSYDYMLVNPVKIVSHSVQAYLPTGEGCLSVDDNVAG 120
QY 121 LVHRNKITIKAKDIEGNDIQLRLKGYPAIVFQHEIDLNGVMFYDHDKHPLOPHPTDA 180
DB 121 LVHRNRITIKAKDIEGNDIQLRLKGYPAIVFQHEIDLNGVMFYDHDKHPLOPHPTDA 180
QY 181 VEV 183
DB 181 VEV 183

RESULT 3
DEF STAAM STANDARD; PRT; 183 AA.
AC Q8CFN4;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Peptide deformylase (EC 3.5.1.88) (PDF) (Polypeptide deformylase).
GN DEF OR SE0789.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RX PubMed=12950922;
RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,
RA Qian Z.-Q., Mao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,
RA Yuan Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.;
RT "Genome-based analysis of virulence genes in a non-biofilm-forming
Staphylococcus epidermidis strain (ATCC 12228).";
RL Mol. Microbiol. 49:1577-1593(2003).
CC -!- FUNCTION: Removes the formyl group from the N-terminal Met of
CC newly synthesized proteins. Requires at least a dipeptide for an
CC efficient rate of reaction. N-terminal L-methionine is a
CC prerequisite for activity but the enzyme has broad specificity at
CC other positions (By similarity).
CC -!- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +
CC methionyl peptide.
CC -!- COFACTOR: Binds 1 iron(II) ion (By similarity).
CC -!- SIMILARITY: Belongs to the polypeptide deformylase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```


QY 61 LAAPQINISKRMIAYLIPDDSGSKSYDMLVNPVKIVSHSVORAYLPTGEGCLSVDDNVAG 120
 DB 61 LAAPQINISKRMIAYLIPDDSGSKSYDMLVNPVKIVSHSVORAYLPTGEGCLSVDDNVAG 119
 QY 121 LVHRNKITIKAKDIEGNDIQRLKGYPAIVFQHEIDLHNGVMFYDHDKDHLPQPHYDA 180
 DB 120 YVPRYARIRVKGTTILEGNDIRLKGFPFAIVFQHEIDLHNGVMFYDHDKDHLPQPHYDA 179
 QY 181 VEVHQ 185
 DB 180 IAIER 184

RESULT 5

DEF2_BACCR
 ID DEF2_BACCR STANDARD; PRT; 184 AA.
 AC Q819K2;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Peptide deformylase 2 (EC 3.5.1.88) (PDF 2) (Polypeptide deformylase
 2)
 GN DEF2 OR BC3974.
 OS Bacillus cereus (strain ATCC 14579 / DSM 31).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=226900;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22608415; PubMed=12721630;
 RA Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B.,
 Kaprat L., Bhattacharya A., Reznik G., Mikhailova N., Lapidus A.,
 Chu L., Mazur M., Goltzman E., Larsen N., D'Souza M., Walunas T.,
 Grechkin Y., Pusch G., Haselkorn R., Fonstein M., Ehrlich S.D.,
 Overbeek R., Kyrpides N.
 RT "Genome sequence of Bacillus cereus and comparative analysis with
 Bacillus anthracis.";
 RL Nature 423:87-91(2003).

CC -!- FUNCTION: Removes the formyl group from the N-terminal Met of
 newly synthesized proteins. Requires at least a dipeptide for an
 efficient rate of reaction. N-terminal L-methionine is a
 prerequisite for activity but the enzyme has broad specificity at
 other positions (By similarity).
 CC -!- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +
 methionyl peptide.

CC -!- COFACTOR: Binds 1 iron(II) ion (By similarity).
 CC -!- SIMILARITY: Belongs to the polypeptide deformylase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).

EMBL; AB017011; AAP10894.1; -
 HAMAP; MF_00163; -; 1.
 DR InterPro; IPR000181; Fmet deformylase.
 DR Pfam; PF01327; Pep deformylase; 1.
 DR PRINTS; PR01576; PDEFORMLASE.
 DR PRODOM; PD003844; Pep deformylase; 1.
 DR TIGRFAMs; TIGR00079; Pept deformyl; 1.
 KW Protein biosynthesis; Hydrolyase; Iron; Complete proteome.
 FT ACT SITE 154 154 BY SIMILARITY.
 FT METAL 110 110 IRON (BY SIMILARITY).
 FT METAL 153 153 IRON (BY SIMILARITY).
 FT METAL 157 157 IRON (BY SIMILARITY).
 SQ SEQUENCE 184 AA; 20474 MW; 854E1CBE1CACA1F CRC64;

Query Match 55.5%; Score 550.5; DB 1; Length 184;
 Best Local Similarity 58.9%; Pred. No. 4.8e-39;
 Matches 106; Conservative 28; Mismatches 45; Indels 1; Gaps 1;

QY 1 MLTKMDIIRDGHPTLRQKALELPLTKKEKETLIAMREFLVNSODEETAKRYGLRSVG 60
 DB 1 MLTKMDIIRDGHPTLRQKALELPLTKKEKETLIAMREFLVNSODEETAKRYGLRSVG 60
 QY 61 LAAPQINISKRMIAYLIPDDSGSKSYDMLVNPVKIVSHSVORAYLPTGEGCLSVDDNVAG 120
 DB 61 LAAPQIGVSKMIAYHV-TDADGTLYSHALFNPKLIHSVVERTYLOGGEGCLSVDRVPG 119
 QY 121 LVHRNKITIKAKDIEGNDIQRLKGYPAIVFQHEIDLHNGVMFYDHDKDHLPQPHYDA 180
 DB 120 YVPRYTRITVKATSIINGEVEVKLRKGLPAIVFQHEIDLHNGVMFYDHDKDHLPQPHYDA 179

RESULT 6

DEF_OCEIH
 ID DEF_OCEIH STANDARD; PRT; 183 AA.
 AC Q8ER96;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Peptide deformylase (EC 3.5.1.88) (PDF) (Polypeptide deformylase).
 GN DEF OR O81410.
 OS Oceanobacillus ihayensis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
 OX NCBI_TaxID=182710;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=HTE831 / DSM 14371 / JCM 11309;
 RX MEDLINE=2220767; PubMed=12235376;
 RA Takami H., Takaki Y., Uchiyama I.;
 RT "Genome sequence of Oceanobacillus ihayensis isolated from the Iheya
 Ridge and its unexpected adaptive capabilities to extreme
 environments.";
 RL Nucleic Acids Res. 30:3927-3935(2002).

CC -!- FUNCTION: Removes the formyl group from the N-terminal Met of
 newly synthesized proteins. Requires at least a dipeptide for an
 efficient rate of reaction. N-terminal L-methionine is a
 prerequisite for activity but the enzyme has broad specificity at
 other positions (By similarity).
 CC -!- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +
 methionyl peptide.

CC -!- COFACTOR: Binds 1 iron(II) ion (By similarity).
 CC -!- SIMILARITY: Belongs to the polypeptide deformylase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).

EMBL; AP004597; BAC13366.1; -
 HAMAP; MF_00163; -; 1.
 DR InterPro; IPR000181; Pep deformylase.
 DR Pfam; PF01327; Pep deformylase; 1.
 DR PRINTS; PR01576; PDEFORMLASE.
 DR PRODOM; PD003844; Pep deformylase; 1.
 DR TIGRFAMs; TIGR00079; Pept deformyl; 1.
 KW Protein biosynthesis; Hydrolyase; Iron; Complete proteome.
 FT ACT SITE 154 154 BY SIMILARITY.
 FT METAL 110 110 IRON (BY SIMILARITY).
 FT METAL 153 153 IRON (BY SIMILARITY).
 FT METAL 157 157 IRON (BY SIMILARITY).
 SQ SEQUENCE 183 AA; 20555 MW; 8582AF4DBE311897 CRC64;

Query Match 55.3%; Score 548.5; DB 1; Length 183;
 Best Local Similarity 60.1%; Pred. No. 7e-39;
 Matches 104; Conservative 34; Mismatches 34; Indels 1; Gaps 1;

QY 1 MLTKMDIIRDGHPTLRQKALELPLTKKEKETLIAMREFLVNSODEETAKRYGLRSVG 60
 DB 1 MLTKMDIIRREGHPSLTRSAWVEVPLSKDKQLLEDMMQFLKNSQDEETAEKRYELRAGV 60

```

Qy 61 LAAPQINISKRMIAVLIIIPDDGSGKSYDMLVNPVKIVSHSVQEAYLPTGGCLSVDDNVAG 120
Db 61 IAAFGLEGIEKIIAHF-EDIDGKLYSMGLVNPXIIHSVSGSYLSSGEGCLSVDRPVEG 119
Qy 121 LVHSHNKTIKAKIDEGNDIQIRLKGYPATVQHEIDHLNGWVFDHDKDHP 173
Db 120 YVPHRARIITKATINDQFVKRLKGYPATVQHEIDHNGIMFDRINTEDP 172

RESULT 7
DEF2_BACAA ID DEF2_BACAA STANDARD; PRT; 184 AA.
AC Q81M09;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Peptide deformylase 2 (EC 3.5.1.88) (PDF 2) (Polypeptide deformylase
DE 2).
DN DEF2 OR BA4187.
OS Bacillus anthracis (strain Ames).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxId=198094;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22608414; PubMed=12721629;
RA Read T.D., Peterson S.N., Tourasse N., Baillie L.W., Paulsen I.T.,
RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
RA Holtzapple E.K., Ostad O.A., Helgason E., Rikstone J., Wu M.,
RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.,
RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
RA Benton J.L., Mahmoud Y., Jiang L., Hance I.R., Weidman J.F.,
RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Niemman W.C.,
RA Hazen A., Cline R., Redmond C., Thwaite J.E., White O., Salzberg S.L.,
RA Thomson B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolsto A.-B.,
RA Fraser C.M.;
RT "The genome sequence of Bacillus anthracis Ames and comparison to
RT closely related bacteria.";
RL Nature 423:81-86(2003).
SC -1- FUNCTION: Removes the formyl group from the N-terminal Met of
SC newly synthesized proteins. Requires at least a dipeptide for an
SC efficient rate of reaction. N-terminal L-methionine is a
SC prerequisite for activity but the enzyme has broad specificity at
SC other positions (By similarity).
SC -1- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +
SC methionyl peptide.
SC -1- COFACTOR: Binds 1 iron(II) ion (By similarity).
SC -1- SIMILARITY: Belongs to the polypeptide deformylase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; AB017037; AAP27909.1; -.
TIGR; BA4187; -.
HAMAP; MF_00163; -.
InterPro; IPR000181; Pmet deformylase.
Pfam; PF01327; Pep deformylase; 1.
PRINTS; PR01576; PDEFORMYLASE.
ProDom; PD003844; Pep deformylase; 1.
TIGRPFAM; TIGR00079; pept deformyl; 1.
Protein biosynthesis; Hydrolase; Iron; Complete proteome.
ACT SITE 154 154 BY SIMILARITY.
METAL 110 110 IRON (BY SIMILARITY).
METAL 153 153 IRON (BY SIMILARITY).
METAL 157 157 IRON (BY SIMILARITY).
SEQUENCE 184 AA; 20515 MW; 12785DF528B0A91F CRC64;

```

```

Query Match          55.3%;   Score 548.5;   DB 1;   Length 184;
Best Local Similarity 60.7%;   Pred. No. 7e-39;
Matches 105;   Conservative 28;   Mismatches 39;   Indels 1;   Gaps 1

Qy 1 MLTWKDIIRDHPTLRQKAALEPLTKKEKETLIAMREFVWNSQDEIAKRYGLRSGVG 60
Db 1 MLTWKDVIREGDPILRNVAEVIIPASEEDNTLTKEMIEFVINSQDEPMASKYSLRPGIG 60

Qy 61 LAAPQINISKRMIAVLIPDDGSKSYDYMLVNPKNVSHSVQEAFLPTGEGCLSDVDNVA 120
Db 61 LAAPQIGISKMIYAVHVTDT-DGTLYSHALFNPKIISHSVERTYLGSGEGCLSDVREVP 119

Qy 121 LVARENKITTAKDIEGNDIQLRKGYPAIVFOHEIDHLNGWMEFYDHDKDHP 173
Db 120 YVRYRITVKAISNGEEVVKLRKGLPAIVFOHEIDHLNGWMEFYDHINKENP 172

RESULT 8
DEF2_BACST
ID DEF2_BACST STANDARD; PRT; 184 AA.
AC Q31410;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Peptide deformylase 2 (EC 3.5.1.88) (PDF 2) (Polypeptide deformylase
DE 2).
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=ATCC 1518;
RX MEDLINE=97272005; PubMed=9126850;
RA Meinel T., Lazemec C., Villoling S., Blanquet S.;
RT "Structure-function relationships within the peptide deformylase
RT family. Evidence for a conserved architecture of the active site
RT involving three conserved motifs and a metal ion.";
RL J. Mol. Biol. 267:749-761(1997).
CC -!- FUNCTION: Removes the formyl group from the N-terminal Met of
CC newly synthesized proteins. Requires at least a dipeptide for an
CC efficient rate of reaction. N-terminal L-methionine is a
CC prerequisite for activity but the enzyme has broad specificity at
CC other positions
CC -!- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +
CC methionyl peptide.
CC -!- COFACTOR: Binds 1 iron(II) ion (By similarity).
CC -!- SIMILARITY: Belongs to the polypeptide deformylase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: Y10549; CAA71591.1; -.
CC FDB: 1LQY; 24-JUL-02.
CC HAMAP: MF_00163; -. 1.
CC InterPro: IPR000181; Pep_deformylase.
CC Pfam: PF01327; Pep_deformylase; 1.
CC PRINTS: PR01576; PDEFORMLYASE.
CC ProDom: PD003844; Pep_deformylase; 1.
CC TIGRFAMs: TIGR00079; pep_deformyl; 1.
CC Protein biosynthesis; Hydrolase; Iron; 3D-structure.
CC ACT SITE 154 154
CC METAL 110 110 IRON (BY SIMILARITY).
CC METAL 153 153 IRON (BY SIMILARITY).
CC METAL 157 157 IRON (BY SIMILARITY).
CC SEQUENCE 184 AA; 20392 MW; 9CD85DEE53632FA0 CRC64;
CC
Qy Query Match          54.5%;   Score 540.5;   DB 1;   Length 184;
Db Best Local Similarity 56.8%;   Pred. No. 3.3e-38;

```

DR	ProDom; PD003844; Pep deformylase; 1.
DR	TIGRFAMs; TIGR00079; pept deformyl; 1.
KW	Protein biosynthesis; Hydrolase; Iron; Complete proteome.
FT	ACT SITE 154 154 BY SIMILARITY.
FT	METAL 110 110 IRON (BY SIMILARITY).
FT	METAL 153 153 IRON (BY SIMILARITY).
FT	METAL 157 157 IRON (BY SIMILARITY).
SQ	SEQUENCE 183 AA; 20610 MW; ID7B2637B2B73D59 CRC64;

Query Match	
Best Local Similarity	54.1%; Score 536.5; DB 1; Length 183;
Matches 106; Conservative 28; Mismatches 44; Indels 3; Gaps 2	

Qy	1	MLTWKDIIRDGHPTLRQAALPLTKKEKETLIAMREFLVNSQDEBEIAKRYGLRSVG	60
Dd	1	MLTWMDDIVREGHPALREVATEVTPFLSDEEKKLGHDMLEFLINSQDEELAERYGLRGVG	60
Qy	61	LAAPOINISKRMTAVLPDDGSKSYDYMLVNPKIVSHSVQEAYLPFGEGCLSDVDNVAG	120
Dd	61	IAPAQVAATKRFIAIHVHDE-KORLSYVLNPKIRSHVQQACLSGGEGCLSDVREVP	119
Qy	121	LVRHNKIITIKAKDIEGNDIQLRLKGYPALVFQFEIDHLNGVMFYDHDKDHP--LOPHT	178
Dd	120	YVVSERVTTDAFDENGTPKLKRFKGYPAIVQVEIDHLNGVMFYDHINKENPSYLPDV	179
Qy	179 D 179		
Dd	180 D 180		

RESULT 10	
DEF_LISMO	ID DEF LISMO STANDARD; PRT; 183 AA.
AC	Q8Y866;
DT	28-FEB-2003 (Rel. 41, Created)
DT	28-FEB-2003 (Rel. 41, Last sequence update)
DT	28-FEB-2003 (Rel. 41, Last annotation update)
DE	Peptide deformylase [EC 3.5.1.88] (PDF) (Polypeptide deformylase).
GN	DEF OR LM01051.
OS	Listeria monocytogenes.
OC	Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX	NCBI_TaxID=1639;
RX	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=EGD-e / Serovar 1/2a;
RX	MEDLINE=21537279; PubMed=11679669;
RA	Glaser P., Frangeul L., Bucherier C., Ruenick C., Amend A.,
RA	Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA	Charbit A., Chetoui F., Couve E., de Daruvar A., Dehoux P.,
RA	Domann B., Dominguez-Bernal G., Duchaud E., Durant L., Duesurget O.,
RA	Ertian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA	Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA	Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapat G.,
RA	Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA	Nordieck G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA	Rommel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA	Vazquez-Boland J.-A., Voss H., Weiland J., Cossart P.;
RT	"Comparative genomics of Listeria species.";
RL	Science 294:849-852(2001).

CC	-!- FUNCTION: Removes the formyl group from the N-terminal Met of newly synthesized proteins. Requires at least a dipeptide for an efficient rate of reaction. N-terminal L-methionine is a prerequisite for activity but the enzyme has broad specificity at other positions (By similarity).
CC	-!- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate + methionyl peptide.
CC	-!- COFACTOR: Binds 1 iron(II) ion (By similarity).
CC	-!- SIMILARITY: Belongs to the polypeptide deformylase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AL591977; CAC99129.1; --
 DR PIR; AC1206; AC1206.
 DR ListiList; LMO01051; --
 DR HAMAP; MF_00163; --; 1.
 DR InterPro; IPR000181; Pep deformylase.
 DR Pfam; PF01327; Pep deformylase; 1.
 DR PRINTS; PR01576; PDEFORMYLASE.
 DR ProDom; PD003844; Pep deformylase; 1.
 DR TIGRFAMs; TIGR00079; pep deformyl; 1.
 KW Protein biosynthesis; Hydrolase; Iron; Complete proteome.
 FT ACT SITE 154 154 BY SIMILARITY.
 FT METAL 110 110 IRON (BY SIMILARITY).
 FT METAL 153 153 IRON (BY SIMILARITY).
 FT METAL 157 157 IRON (BY SIMILARITY).
 SQ SEQUENCE 183 AA; 20643 MW; 6582430603CDA4EF CRC64;

Query Match 53.8%; Score 533.5; DB 1; Length 183;
 Best Local Similarity 57.5%; Pred. No. 1.2e-37;
 Matches 104; Conservative 30; Mismatches 44; Indels 3; Gaps 2;

QY 1 MLTMDKIIRDGHPTLRQKAAELPLTKKEETLIAMREFLVNSQDEIAKRYGLRSGVG 60
 Db 1 MLTMDIIVREGHPALREVATEVTPPLSDDEKKLGRDMLFELINSQDEIAKRYGLRSGVG 60
 QY 61 LAAPQINISKEMIAVLIPDDGSGKSYDMLNPKIVSHSVQEAFLPTGEGCLSDVDNVAG 120
 Db 61 IAAQLAVTKKFLAHVHDE-KDRLYSVLVNPKIRSHSVQOACLSGEGCLSDVDREVP 119
 QY 121 LVHRHNKTIKAKDIEGNDIQLRLKGYPAIVFQHEIDHLNGWVFYDHDKDP--LQPH 178
 Db 120 YVRSERVITDAFDENGTPKLRFKGYPAIVFQHEIDHLNGWVFYDHDKDP--LQPH 178
 QY 179 D 179
 Db 180 D 180

RESULT 11
 DEF_BACHD STANDARD; PRT; 182 AA.
 AC Q9K19; 2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Peptide deformylase (EC 3.5.1.88) (PDF) (Polypeptide deformylase).
 GN DEF OR BH2658.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis."
 RL Nucleic Acids Res. 28:4317-4331(2000).
 CC -!- FUNCTION: Removes the formyl group from the N-terminal Met of
 CC newly synthesized proteins. Requires at least a dipeptide for an
 CC efficient rate of reaction. N-terminal L-methionine is a
 CC prerequisite for activity but the enzyme has broad specificity at
 CC other positions (By similarity).
 CC -!- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +
 CC methionyl peptide.
 CC -!- COFACTOR: Binds 1 iron(II) ion (By similarity).
 CC -!- SIMILARITY: Belongs to the polypeptide deformylase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AP001516; BAB06377.1; --
 DR PIR; B83982; B83982.
 DR HSP; P27251; 2DEF.
 DR HAMAP; MF_00163; --; 1.
 DR InterPro; IPR000181; Pep deformylase.
 DR Pfam; PF01327; Pep deformylase; 1.
 DR PRINTS; PR01576; PDEFORMYLASE.
 DR ProDom; PD003844; Pep deformylase; 1.
 DR TIGRFAMs; TIGR00079; pep deformyl; 1.
 KW Protein biosynthesis; Hydrolase; Iron; Complete proteome.
 FT ACT SITE 154 154 BY SIMILARITY.
 FT METAL 110 110 IRON (BY SIMILARITY).
 FT METAL 153 153 IRON (BY SIMILARITY).
 FT METAL 157 157 IRON (BY SIMILARITY).
 SQ SEQUENCE 182 AA; 20599 MW; 665D39B56EB6153F CRC64;

Query Match 51.1%; Score 506.5; DB 1; Length 182;
 Best Local Similarity 58.0%; Pred. No. 2.2e-35;
 Matches 98; Conservative 31; Mismatches 39; Indels 1; Gaps 1;

QY 1 MLTMDKIIRDGHPTLRQKAAELPLTKKEETLIAMREFLVNSQDEIAKRYGLRSGVG 60
 Db 1 MLTMDIIVREGHPALREVATEVTPPLSDDEKKLGRDMLFELINSQDEIAKRYGLRSGVG 60
 QY 61 LAAPQINISKEMIAVLIPDDGSGKSYDMLNPKIVSHSVQEAFLPTGEGCLSDVDNVAG 120
 Db 61 LAAPQIGLSQMIATVHTDENE-KEYSLVLFNPKIISSEVMTLEGEGCLSDVDREVOG 119
 QY 121 LVHRHNKTIKAKDIEGNDIQLRLKGYPAIVFQHEIDHLNGWVFYDHDKDP--LQPH 169
 Db 120 IVPFHARITVKAINENNEVRLLKGFPAIVFQHEIDHLNGWVFYDHDKDP--LQPH 168

RESULT 12
 DEF_LACPL STANDARD; PRT; 186 AA.
 AC Q88VB2;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Peptide deformylase (EC 3.5.1.88) (PDF) (Polypeptide deformylase).
 GN DEF OR DEF1 OR LP 2155.
 OS Lactobacillus plantarum.
 OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
 OC Lactobacillus.
 OX NCBI_TaxID=1590;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCIMB 8826 / WCFS1;
 RX MEDLINE=22480296; PubMed=12566566;
 RA Kleerebezem M., Boekhorst J., van Kranenburg R., Molenaar D.,
 RA Kuipers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,
 RA Fiers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,
 RA Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,
 RA De Vos W.M., Siezen R.J.;
 RT "Complete genome sequence of Lactobacillus plantarum WCFS1."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).
 CC -!- FUNCTION: Removes the formyl group from the N-terminal Met of
 CC newly synthesized proteins. Requires at least a dipeptide for an
 CC efficient rate of reaction. N-terminal L-methionine is a
 CC prerequisite for activity but the enzyme has broad specificity at
 CC other positions (By similarity).
 CC -!- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +
 CC methionyl peptide.
 CC -!- COFACTOR: Binds 1 iron(II) ion (By similarity).


```

CC CC -!- SIMILARITY: Belongs to the polypeptide deformylase family.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; AL935258; CAD64511.1; --
CC CC HAMAP; MF 00163; --; 1.
CC CC InterPro; IPR000181; Pep deformylase.
CC CC Pfam; PF01327; Pep deformylase; 1.
CC CC PRINTS; PD01576; PDEFORMYLASE.
CC CC Protein biosynthesis; Hydrolase; Iron; Complete proteome.
CC CC ACT_SITE 157 157 BY SIMILARITY.
CC CC METAL 113 113 IRON (BY SIMILARITY).
CC CC METAL 156 156 IRON (BY SIMILARITY).
CC CC METAL 160 160 IRON (BY SIMILARITY).
CC CC SEQUENCE 186 AA; 20854 MW; E080FF56D7723576 CRC64;
CC CC -----
CC CC Query Match 49.2%; Score 488; DB 1; Length 186;
CC CC Best Local Similarity 49.7%; Pred. No. 7.9e-34;
CC CC Matches 91; Conservative 45; Mismatches 45; Indels 2; Gaps 1;
CC CC -----
CC CC QY 1 MLTWKDIIRDGHTLROKAAELEPLTKKEKTLIAMEFVLNSQDEEIAKRYGLRSGVG 60
CC CC DB 1 MIKMDIIRREGNHTLRAEAKQVFFLSEADQKLANDMWEYLENSQDPELAKYGLRAGVG 60
CC CC QY 61 LAAPQINISKRMIAVLIPDDSGSKS--YDMLVNPVKIVSHSVQEAAYLPTGEGCLSVDDNV 118
CC CC DB 61 LAAPQVDVSEQAAVLVPSSEDDPEVFKDVIINPVIISHSVQPGALTEGEGCLSVDRDI 120
CC CC QY 119 AGLVHRHNKTIKAKDIEGNDIQLRLKGYPAIVFOHEIDLHNGVMFYDHIDKDPHQPT 178
CC CC DB 121 AGYVIRHDIILRYNNMAGEBKRLKNYPALVCOHEIDLHLGLIUFYDHINGDNPFAADD 180
CC CC QY 179 DAV 181
CC CC DB 181 DLV 183
CC CC -----
CC CC RESULT 13
CC CC DEF_ENTFA STANDARD; PRT; 187 AA.
CC CC ID DEF_ENTFA Q822J0;
CC CC AC Q822J0;
CC CC DT 15-MAR-2004 (Rel. 43, Created)
CC CC DT 15-MAR-2004 (Rel. 43, Last sequence update)
CC CC DT 15-MAR-2004 (Rel. 43, Last annotation update)
CC CC DE Peptide deformylase (EC 3.5.1.88) (PDF) (Polypeptide deformylase).
CC CC GN DEF OR EF3066.
CC CC OS Enterococcus faecalis (Streptococcus faecalis).
CC CC OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
CC CC OX NCBI_TaxID=1351;
CC CC RN [1]
CC CC RP SEQUENCE FROM N.A.
CC CC RC STRAIN=V583 / ATCC 700802;
CC CC RX MEDLINE=22550857; PubMed=12663927;
CC CC RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
CC CC Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
CC CC Tettelin H., Dodson R.J., Umayam L., Brinkay L., Beanan M.,
CC CC Daugherty S., DeBoy R.T., Durkin S., Kolonay J., Madupu R., Nelson W.,
CC CC Vamathevan J., Tran B., Upton J., Hansen T., Shetty J., Khouri H.,
CC CC Utterback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.;
CC CC RT "Role of mobile DNA in the evolution of vancomycin-resistant
CC CC Enterococcus faecalis";
CC CC RL Science 299:2071-2074(2003).
CC CC -!- FUNCTION: Removes the formyl group from the N-terminal Met of
CC CC newly synthesized proteins. Requires at least a dipeptide for an
CC CC efficient rate of reaction. N-terminal L-methionine is a
CC CC prerequisite for activity but the enzyme has broad specificity at
CC CC other positions (By similarity).

```

```

CC CC -!- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +
CC CC methionyl peptide.
CC CC -!- COFACTOR: Binds 1 iron(II) ion (By similarity).
CC CC -!- SIMILARITY: Belongs to the polypeptide deformylase family.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; AE016956; AAO82749.1; --
CC CC TIGR; EF3066; --; 1.
CC CC HAMAP; MF 00163; --; 1.
CC CC InterPro; IPR000181; Fmet deformylase.
CC CC Pfam; PF01327; Pep deformylase; 1.
CC CC PRINTS; PD01576; PDEFORMYLASE.
CC CC ProDom; PD03844; Pep deformylase; 1.
CC CC TIGRFAMs; TIGR00079; Pep deformyl; 1.
CC CC Protein biosynthesis; Hydrolase; Iron; Complete proteome.
CC CC ACT_SITE 158 158 BY SIMILARITY.
CC CC METAL 114 114 IRON (BY SIMILARITY).
CC CC METAL 157 157 IRON (BY SIMILARITY).
CC CC METAL 161 161 IRON (BY SIMILARITY).
CC CC SEQUENCE 187 AA; 20912 MW; 9CAF46335311B0E2 CRC64;
CC CC -----
CC CC Query Match 48.3%; Score 478.5; DB 1; Length 187;
CC CC Best Local Similarity 55.1%; Pred. No. 4.9e-33;
CC CC Matches 97; Conservative 30; Mismatches 46; Indels 3; Gaps 1;
CC CC -----
CC CC QY 1 MLTWKDIIRDGHTLROKAAELEPLTKKEKTLIAMEFVLNSQDEEIAKRYGLRSGVG 60
CC CC DB 1 MITWKDIIRREGNHTLRVAEVPVITEEDRQGLDMLTLKNSQDPVKAELOLRGGVG 60
CC CC QY 61 LAAPQINISKRMIAVLIPD---DGSKGSDYMLVNPVKIVSHSVQEAAYLPTGEGCLSVDDN 117
CC CC DB 61 LAAPQDISKRIIAVHVPSNDPENETPSLTVMNPKILSHSVQDVCLEGGCLSVDRD 120
CC CC QY 118 VAGLVHRHNKTIKAKDIEGNDIQLRLKGYPAIVFOHEIDLHNGVMFYDHIDKDPH 173
CC CC DB 121 VPGYVVRHNKTIIVSYFDWAGEKHVKRLKNYEAIVQHEIDHNGIMFYDHINKENP 176
CC CC -----
CC CC RESULT 14
CC CC DEF_STRPY STANDARD; PRT; 204 AA.
CC CC ID DEF_STRPY Q99XY7; P82590;
CC CC AC Q99XY7; P82590;
CC CC DT 28-FEB-2003 (Rel. 41, Created)
CC CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC CC DT 15-MAR-2004 (Rel. 43, Last annotation update)
CC CC DE Peptide deformylase (EC 3.5.1.88) (PDF) (Polypeptide deformylase).
CC CC GN DEF OR SPY198 OR SPYM3_1684 OR SP51686.
CC CC OS Streptococcus pyogenes, and
CC CC OC Streptococcus pyogenes (serotype M3).
CC CC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
CC CC Streptococcus.
CC CC OX NCBI_TaxID=1314, 198466;
CC CC RN [1]
CC CC RP SEQUENCE FROM N.A.
CC CC RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
CC CC RX MEDLINE=21192684; PubMed=11296296;
CC CC RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savić G., Lyon K.,
CC CC Primeaux C., Szatog S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
CC CC Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
CC CC Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
CC CC RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes";
CC CC Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
CC CC [2]
CC CC RL SEQUENCE FROM N.A.
CC CC RP STRAIN=MGAS315 / Serotype M3;
CC CC RX MEDLINE=22133808; PubMed=12122206;

```



```

QY 1 MLTMKDIIIRDGHPTLRQAAELEPLTKKEKETLIAMREFLVNSODEETAKRYGLRSGVG 60
Db 13 LITMDIIIREGNPTLRVAKESVSLPLCDEIILGKMWQFLKHSQDPVMAEKLGLRAGVG 72
QY 61 LAAPOINISKRMIAVLIPD--DGSCK-----SYDMLVNPKNVSHSVQBYLPTGEGCL 112
Db 73 LAAPOIDVSKRIIAVLVPLPDKEGNPPKEAYSQOEVLNPNKIVSHSVQDAALSDGEGCL 132
QY 113 SVDDNVAGLVHRHNKITTAKDIEGNDIQLRLKGYPAIVFOHEIDHNLGVMEFYDHIDKH 172
Db 133 SVDRVVEGVVVEHARVTVDYDKESQOHRKIKLKGYNALVQHEIDHNGILFYDRINAKN 192
QY 173 PLQ 175
Db 193 PFE 195

```

Search completed: March 31, 2004, 16:35:32
Job time : 19 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 16:30:30 ; Search time 20 Seconds
(without alignments)
909.010 Million cell updates/sec

Title: US-09-896-580B-12
Perfect score: 991
Sequence: 1 MLTMKDIIRDGHPTLRQKAA.....KDHPHQPTDAVEVHQHHH 189
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	938	94.7	183	2 A89879	hypothetical prote
2	575.5	58.1	184	1 D69862	formylmethionine d
3	536.5	54.1	183	2 AB1563	formylmethionine d
4	533.5	53.8	183	2 AC1206	formylmethionine d
5	506.5	51.1	182	2 B83982	formylmethionine d
6	451.5	45.6	203	2 B98035	formylmethionine d
7	450.5	45.5	203	2 B95169	polypeptide deform
8	426	43.0	211	2 H86694	polypeptide deform
9	273.5	27.6	198	2 F82886	polypeptide deform
10	263	26.5	198	2 A90598	hypothetical prote
11	228.5	23.1	216	2 S73913	polypeptide deform
12	222.5	22.5	224	2 H97709	formylmethionine d
13	221.5	22.4	226	1 G64211	formylmethionine d
14	218.5	22.0	168	2 C82494	polypeptide deform
15	217.5	21.9	179	2 A12056	polypeptide deform
16	215.5	21.7	191	2 A55228	fms protein homolo
17	211.5	21.3	169	1 D64082	N-formylmethionyla
18	207	20.9	181	2 C81580	polypeptide deform
19	205.5	20.7	232	2 H75274	polypeptide deform
20	203.5	20.5	273	2 T48639	hypothetical prote
21	203	20.5	181	2 B71526	probable polypepti
22	198	20.0	164	2 C72224	polypeptide deform
23	193.5	19.5	170	2 AF2621	polypeptide deform
24	193.5	19.5	170	2 F97403	polypeptide deform
25	190	19.2	173	2 E84987	polypeptide deform
26	189	19.1	150	2 A97205	formylmethionine d
27	184.5	18.6	160	1 F69613	N-formylmethionyl-
28	184.5	18.6	169	2 AB1010	polypeptide deform
29	183	18.5	259	2 B86288	hypothetical prote

polypeptide deform
formylmethionine d
N-formylmethionyla
peptide deformylas
peptide deformylas
polypeptide deform
N-formylmethionyl-
formylmethionine d
polypeptide deform
formylmethionine d
polypeptide deform
polypeptide deform
polypeptide deform
polypeptide deform
polypeptide deform
polypeptide deform

ALIGNMENTS

RESULT 1

A89879
hypothetical protein pdf1 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 31-Dec-2001
C:Accession: A89879
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: A89879
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-183 <KUR>
A:Cross-references: GB:BA000018; PID:g13700892; PIDN:BAB42188.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: pdf1
C:Superfamily: polypeptide deformylase

Query Match	94.7%	Score	938	DB 2	Length	183
Best Local Similarity	98.9%	Pred. No.	5e-72			
Matches	181	Conservative	2	Mismatches	0	Gaps
QY	1	MLTMKDIIRDGHPTLRQKAAELPLTKKEKFTLIAMREFLVNSQDEETAKRYGLRSGVG	60			
Db	1	MLTMKDIIRDGHPTLRQKAAELPLTKKEKFTLIAMREFLVNSQDEETAKRYGLRSGVG	60			
QY	61	LAAPQINIKRMIAVLIPDDGSGKSYDYMVNPKIVSHSVQEAFLPTGEGCLSVDDNVAG	120			
Db	61	LAAPQINIKRMIAVLIPDDGSGKSYDYMVNPKIVSHSVQEAFLPTGEGCLSVDDNVAG	120			
QY	121	LVRHNRITIKAKDIEGNDIQLRLKGYPAIVFQHEIDHLNGVMFYDHIDKHPLQHTDA	180			
Db	121	LVRHNRITIKAKDIEGNDIQLRLKGYPAIVFQHEIDHLNGVMFYDHIDKHPLQHTDA	180			
QY	181	VEV	183			
Db	181	VEV	183			

RESULT 2

D69862
formylmethionine deformylase homolog ykrB - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C:Accession: D69862
R:Kunat, F.; Ogasawara, N.; Moszer, I.; Alvertini, A.M.; Alloni, G.; Azevedo, V.; Berter, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chc, A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrar, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mausel, Y.M.; Ogawa, K.; Ogiwara, A.; Oudege, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, R.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, A.; Authors: Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: D69862
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-184 <KUN>
A:Cross-references: GB:299111; GB:AL009126; NID:g2633699; PIDN:CAB13329.1; PID:g2633827
A:Experimental source: strain 168
C:Genetics:
A:Gene: ykrB
C:Superfamily: polypeptide deformylase

Query Match 58.1%; Score 575.5; DB 1; Length 184;

Best Local Similarity 59.5%; Pred. No. 28-41;

Matches 110; Conservative 33; Mismatches 41; Indels 1; Gaps 1;

QY 1 MLTMDIIRDGHTPLRQKAAELPLTKKEETLIAMREFLVNSQDEIEAKRYGLRSGVG 60

DB 1 MITMENIVRDGHFALRETAPEVLPPTDAEQQLADMIFFVKNSQNPQLAEKYLKRPVG 60

QY 61 LAAPQINISKRMIAVLIPDDGSGKSYDMLVNPVKIVSHSVQAYLPTGEGCLSVDDNVAG 120

DB 61 LAAPQINISKRMIAV-HAEDASGKLYSALFNPKIVSHSVKSYLTSGEGCLSVDEAIPG 119

QY 121 LVHRENKITTAKDIEGNDIQLRLKGYPAIVFOHEIDLHNGVMFYDHDKDHPLQPHDA 180

DB 120 YVPRVIRVKGTTLEGENDIRLKGFPALVFOHEIDLHNGVMFYDHDKDHKPNFKEPENA 179

QY 181 VEVHQ 185

DB 180 IAIER 184

RESULT 3

AB1563

formylmethionine deformylase and to *B. subtilis* YkrB protein homolog lin1043 [imported]

C:Species: *Listeria innocua*

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C:Accession: AB1563

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.

Science 294, 849-852, 2001

A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma

ok, C.; Schluerer, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,

A:Title: Comparative genomics of *Listeria* species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AB1563

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-183 <GLA>

A:Cross-references: GB:AL592022; PIDN:CAC96274.1; PID:g16413502; GSPDB:GN00178

A:Experimental source: strain Clip11262

C:Genetics:

A:Gene: lin1043

C:Superfamily: polypeptide deformylase

Query Match 54.1%; Score 536.5; DB 2; Length 183;

Best Local Similarity 58.6%; Pred. No. 4e-38;

Matches 106; Conservative 28; Mismatches 44; Indels 3; Gaps 2;

QY 1 MLTMDIIRDGHTPLRQKAAELPLTKKEETLIAMREFLVNSQDEIEAKRYGLRSGVG 60

DB 1 MLTMDIIVREGHPALREVAETVTFPLSDEEKLGHMDLFLNSQDEIEAKRYGLRSGVG 60

QY 61 LAAPQINISKRMIAVLIPDDGSGKSYDMLVNPVKIVSHSVQAYLPTGEGCLSVDDNVAG 120

DB 61 IAAPQLAVTKRFLAIHVHDE-KDRLSYVLYNPKIRSHSVQACLSGGEGCLSVDRVPG 119

QY 121 LVHRENKITTAKDIEGNDIQLRLKGYPAIVFOHEIDLHNGVMFYDHDKDHPLQPHDA 178

DB 120 YVRSERVITDAFDENGTPKLRFKGYPAIVVQHEIDLHNGVMFYDHDKDHKPNFKEPENA 179

QY 179 D 179

DB 180 D 180

RESULT 4

AC1206

formylmethionine deformylase and to *B. subtilis* YkrB protein homolog lmol051 [imported]

C:Species: *Listeria monocytogenes*

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C:Accession: AC1206

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.

Science 294, 849-852, 2001

A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma

ok, C.; Schluerer, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,

A:Title: Comparative genomics of *Listeria* species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AC1206

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-183 <GLA>

A:Cross-references: GB:NC_003210; PIDN:CAC99129.1; PID:g16410453; GSPDB:GN00177

A:Experimental source: strain EGD-e

C:Genetics:

A:Gene: lmol051

C:Superfamily: polypeptide deformylase

Query Match 53.8%; Score 533.5; DB 2; Length 183;

Best Local Similarity 57.5%; Pred. No. 7.1e-38;

Matches 104; Conservative 30; Mismatches 44; Indels 3; Gaps 2;

QY 1 MLTMDIIRDGHTPLRQKAAELPLTKKEETLIAMREFLVNSQDEIEAKRYGLRSGVG 60

DB 1 MLTMDIIVREGHPALREVAETVTFPLSDEEKLGRDMLFLNSQDEIEAKRYGLRSGVG 60

QY 61 LAAPQINISKRMIAVLIPDDGSGKSYDMLVNPVKIVSHSVQAYLPTGEGCLSVDDNVAG 120

DB 61 IAAPQLAVTKRFLAIHVHDE-KDRLSYVLYNPKIRSHSVQACLSGGEGCLSVDRVPG 119

QY 121 LVHRENKITTAKDIEGNDIQLRLKGYPAIVFOHEIDLHNGVMFYDHDKDHPLQPHDA 178

DB 120 YVRSERVITDAFDENGTPKLRFKGYPAIVVQHEIDLHNGVMFYDHDKDHKPNFKEPENA 179

QY 179 D 179

DB 180 D 180

RESULT 5

B83982

formylmethionine deformylase BH2658 [imported] - *Bacillus halodurans* (strain C-125)

C:Species: *Bacillus halodurans*

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C:Accession: B83982

R:Itakami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hir

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: B83982

A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-182 <STO>
 A:Cross-references: GB:AP001516; GB:BA000004; NID:g10175192; PIDN:BA06377.1; GSPDB:GN00164
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: BH2658
 C:Superfamily: polypeptide deformylase

Query Match 51.1%; Score 506.5; DB 2; Length 182;
 Best Local Similarity 58.0%; Pred. No. 1.3e-35;
 Matches 98; Conservative 31; Mismatches 39; Indels 1; Gaps 1;

QY 1 MLTMKDIIIRDGHTLRQKAAELEPLTKKEETLIAMREFLVNSQDEETAKRYGLRSGVG 60
 DB 1 MLTMKDIIIRDGHTLRQKAAELEPLTKKEETLIAMREFLVNSQDEETAKRYGLRSGVG 60
 QY 61 LAAPQINISKRMIIVLIPDDGSKSYDY--MLVNPKIYSHSVQEAFLPTGEGCLSVDDNVAG 120
 DB 61 LAAPQINISKRMIIVLIPDDGSKSYDY--MLVNPKIYSHSVQEAFLPTGEGCLSVDDNVAG 120
 QY 121 LVHRHNKTIKADIEGNDIQLRLKGYPAIVFOHEIDLHNGVMFYDHD 169
 DB 120 IVPRHARTVXALNENNEVRLKAGFPPIVFOHEIDLHNGVMFYDRIE 168

RESULT 6
 E98035
 formylmethionine deformylase (EC 3.5.1.31) fms [imported] - Streptococcus pneumoniae (strain TIGR4)
 C:Species: Streptococcus pneumoniae
 C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
 C:Accession: E98035
 R:Hoekings, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; Eick, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Mateushima, P.; McAhren, S.; Meyer, P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001
 A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.; A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
 A:Reference number: A97872; MUID:21429245; PMID:11544234
 A:Accession: E98035
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-203 <KUR>
 A:Cross-references: GB:AE007317; PIDN:AAL00114.1; PID:g15458954; GSPDB:GN00174
 C:Genetics:
 A:Gene: fms
 C:Keywords: hydrolase

Query Match 45.6%; Score 451.5; DB 2; Length 203;
 Best Local Similarity 51.7%; Pred. No. 6.8e-31;
 Matches 93; Conservative 35; Mismatches 45; Indels 7; Gaps 2;

QY 1 MLTMKDIIIRDGHTLRQKAAELEPLTKKEETLIAMREFLVNSQDEETAKRYGLRSGVG 60
 DB 13 LIDMNDIIRGNPTLRVAEEVTPFLSDQEIILGKMMQFLKHSQDPVMAEKMLRGGVG 72
 QY 61 LAAPQINISKRMIIVLIPDDGSKSYDY--MLVNPKIYSHSVQEAFLPTGEGCL 113
 DB 73 LAAPQINISKRMIIVLIPDDGSKSYDY--MLVNPKIYSHSVQEAFLPTGEGCL 113
 QY 114 VDDNVAGLVHRHNKTIKADIEGNDIQLRLKGYPAIVFOHEIDLHNGVMFYDHD 173
 DB 133 VDRNPGYVVRHARVTVDFDKGKHRIKLGKNSIVVQHEIDHNGIMFYDRINEKDP 192

RESULT 7
 E95169
 polypeptide deformylase [imported] - Streptococcus pneumoniae (strain TIGR4)
 C:Species: Streptococcus pneumoniae
 C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
 C:Accession: E95169
 R:Tettelin, H.; Nelson, K.B.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heidrich, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, M.; Nelson, T.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001
 A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison, A.; Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
 A:Reference number: A95000; MUID:21357209; PMID:11463916
 A:Accession: E95169
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-203 <KUR>
 A:Cross-references: GB:AE005672; PIDN:AAK75550.1; PID:g14972945; GSPDB:GN00164; TIGR:SP
 A:Experimental source: strain TIGR4
 C:Genetics:
 A:Gene: SPI456

Query Match 45.5%; Score 450.5; DB 2; Length 203;
 Best Local Similarity 51.7%; Pred. No. 8.3e-31;
 Matches 93; Conservative 35; Mismatches 45; Indels 7; Gaps 2;

QY 1 MLTMKDIIIRDGHTLRQKAAELEPLTKKEETLIAMREFLVNSQDEETAKRYGLRSGVG 60
 DB 13 LIDMNDIIRGNPTLRVAEEVTPFLSDQEIILGKMMQFLKHSQDPVMAEKMLRGGVG 72
 QY 61 LAAPQINISKRMIIVLIPDDGSKSYDY--MLVNPKIYSHSVQEAFLPTGEGCL 113
 DB 73 LAAPQINISKRMIIVLIPDDGSKSYDY--MLVNPKIYSHSVQEAFLPTGEGCL 113
 QY 114 VDDNVAGLVHRHNKTIKADIEGNDIQLRLKGYPAIVFOHEIDLHNGVMFYDHD 173
 DB 133 VDRNPGYVVRHARVTVDFDKGKHRIKLGKNSIVVQHEIDHNGIMFYDRINEKDP 192

RESULT 8
 H86694
 polypeptide deformylase [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
 C:Species: Lactococcus lactis subsp. lactis
 C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
 C:Accession: H86694
 R:Bolotin, A.; Winkler, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich, S.; Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
 A:Reference number: A86625; MUID:21235186; PMID:11337471
 A:Accession: H86694
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-211 <STO>
 A:Cross-references: GB:AE005176; PID:g12723447; PIDN:AAK04658.1; GSPDB:GN00146
 A:Experimental source: strain IL1403
 C:Genetics:
 A:Gene: def

Query Match 43.0%; Score 426; DB 2; Length 211;
 Best Local Similarity 49.2%; Pred. No. 1e-28;
 Matches 91; Conservative 36; Mismatches 46; Indels 12; Gaps 3;

QY 1 MLTMKDIIIRDGHTLRQKAAELEPLTKKEETLIAMREFLVNSQDEETAKRYGLRSGVG 60
 DB 16 MISMDIIRREGYPTLRVANDVTLPLSDQEIILGKMMQFLKHSQDPVMAEKMLRGGVG 75
 QY 61 LAAPQINISKRMIIVLIPDDGSKSYDY--MLVNPKIYSHSVQEAFLPTG 108
 DB 76 LAANQLGLLKKVIAVLIPNEPEVDEGNEIPPEKAYKMEIMYNAKVSHSVQDAVEGG 135
 QY 109 ECLSVDDNVAGLVHRHNKTIKADIEGNDIQLRLKGYPAIVFOHEIDLHNGVMFYDHI 168
 DB 136 ECLSVDRVPGYVVRHARVTVDFDKGKHRIKLGKNSIVVQHEIDHNGIMFYDHI 195

QY 169 DKDHP 173
 DB 196 NWNDP 200

RESULT 9
 F82886
 polypeptide deformylase U0465 [imported] - Ureaplasma urealyticum

Gene: VCA0150
Map position: 2
Superfamily: polypeptide deformylase

Search completed: March 31, 2004, 16:36:57
Job time : 21 secs

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 16:29:29 ; Search time 39 Seconds
(without alignments)
1529.050 Million cell updates/sec

Title: US-09-896-580B-12

Perfect score: 991

Sequence: 1 MLTMKDIIIRDGHPTIRQKAA.....KDHLPHQTDAVEVHQHHH 189

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:.*
1: sp_archaea:.*
2: sp_bacteria:.*
3: sp_fungi:.*
4: sp_human:.*
5: sp_invertebrate:.*
6: sp_mammal:.*
7: sp_mmc:.*
8: sp_organelle:.*
9: sp_phage:.*
10: sp_plant:.*
11: sp_rodent:.*
12: sp_virus:.*
13: sp_vertebrate:.*
14: sp_unclassified:.*
15: sp_rvirus:.*
16: sp_bacteriapi:.*
17: sp_archaea:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	943	95.2	183	Q8NX78	Q8NX78 staphylococ
2	550.5	55.5	184	Q819K2	Q819K2 bacillus ce
3	548.5	55.3	184	Q81MQ9	Q81MQ9 bacillus an
4	488	49.2	186	Q88VB2	Q88VB2 lactobacill
5	478.5	48.3	187	Q822J0	Q822J0 enterococcu
6	472.5	47.7	187	Q842S4	Q842S4 enterococcu
7	453	45.7	204	Q8E378	Q8E378 streptococc
8	453	45.7	204	Q8DXF6	Q8DXF6 streptococc
9	451.5	45.6	203	Q939R9	Q939R9 streptococc
10	451.5	45.6	203	Q8DP79	Q8DP79 streptococc
11	435	43.9	204	Q8DWC2	Q8DWC2 streptococc
12	268.5	27.1	204	Q9KX63	Q9KX63 mycoplasma
13	237.5	24.0	315	Q7XYP8	Q7XYP8 chlorarachn
14	218.5	22.0	166	Q8GDO9	Q8GDO9 heliobacill
15	203.5	20.5	273	Q8LEH0	Q8LEH0 arabidopsis
16	203.5	20.5	273	Q949U8	Q949U8 arabidopsis

17	202	20.4	181	16	Q7VINS	Q7vins helicobacte
18	202	20.4	201	16	Q7UHZ5	Q7uhz5 rhodopirell
19	201	20.3	171	16	Q7VKK9	Q7vkk9 haemophilus
20	199	20.1	192	16	Q7VKG6	Q7vkg6 prochloroco
21	199	20.1	284	10	Q94CZ4	Q94cz4 oryza sativ
22	196.5	19.8	184	16	Q8AAP4	Q8aap4 bacteroides
23	196	19.8	170	16	Q8PGZ0	Q8pgz0 xanthomonas
24	191.5	19.3	177	2	Q8REQ2	Q8req2 zymomonas m
25	191.5	19.3	241	5	Q8I372	Q8i372 plasmodium
26	190.5	19.2	150	16	Q895Q2	Q895q2 clostridium
27	182.5	18.4	169	16	Q83P21	Q83p21 shigella fl
28	180	18.2	185	16	Q82TC8	Q82tc8 nitrosomona
29	179	18.1	170	16	Q8P4F9	Q8p4f9 xanthomonas
30	175	17.7	156	16	Q819U0	Q819u0 bacillus ce
31	175	17.7	201	16	Q7V5F9	Q7v5f9 prochloroco
32	171.5	17.3	176	16	Q82TW4	Q82tw4 nitrosomona
33	169	17.1	201	16	Q7U9D4	Q7u9d4 synecococc
34	168.5	17.0	193	16	Q8FMD0	Q8fmd0 corynebacte
35	168	17.0	156	16	Q81WH1	Q81wh1 bacillus an
36	168	17.0	169	16	Q8FT51	Q8ft51 corynebacte
37	166	16.8	188	16	Q8DIB4	Q8dib4 synecococc
38	164.5	16.6	201	16	Q7V3K7	Q7v3k7 prochloroco
39	164	16.5	196	5	Q8INL3	Q8inl3 drosophila
40	163	16.4	209	16	Q83AK6	Q83ak6 coxiella bu
41	163	16.4	238	5	Q9VGX2	Q9vgx2 drosophila
42	161	16.2	170	16	Q7WOS9	Q7wos9 bordetella
43	161	16.2	170	16	Q7WIV3	Q7wiv3 bordetella
44	161	16.2	170	16	Q7VS88	Q7vs88 bordetella
45	160.5	16.2	193	16	Q8NM41	Q8nm41 corynebacte

ALIGNMENTS

RESULT 1

Q8NX78 ID Q8NX78 PRELIMINARY; PRT; 183 AA.
AC Q8NX78;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Pdf1 protein.
DE PDF1 OR MW0974.
GN Staphylococcus aureus (strain MW2)
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-acquired MRSA";
RL Lancet 359:1819-1827(2002).
DR EMBL; AF004825; BAB94839.1;
DR GO; GO:0008463; F:formylmethionine deformylase activity; IEA.
DR GO; GO:0008412; P:protein biosynthesis; IEA.
DR InterPro; IPR000181; P:protein biosynthesis; IEA.
DR Pfam; PF01327; P:pep deformylase; 1.
DR PRINTS; PR01576; P:DEFORMYLASE.
DR ProDom; PD003844; P:pep deformylase; 1.
DR TIGRFAMs; TIGR00079; P:pep deformylase; 1.
KW Complete proteome.
SQ SEQUENCE 183 AA; 20559 MW; 32A64066AE5CAB0E CRC64;

Query Match 95.2%; Score 943; DB 16; Length 183;
Best Local Similarity 99.5%; Pred No. 4.1e-76;
Matches 182; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLTMKDIIIRDGHPTIRQKAAELEPLTKKEETLIAMREFLVNSQDEETAKYGLRSVG 60
|||||

Db 1 MLTMKDIIIRDGHPTIRQKAAELEPLTKKEETLIAMREFLVNSQDEETAKYGLRSVG 60
|||||

RESULT 3
Q81M09
Q81M09
PRELIMINARY; PRT; 184 AA.
AC Q81M09;
DT 01-JUN-2003 (TREMBlurel. 24, Created)
DT 01-JUN-2003 (TREMBlurel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlurel. 25, Last annotation update)

```

RESULT 4
Q88VB2
ID Q88VB2 PRELIMINARY; PRT; 186 AA.
AC Q88VB2;
DT 01-JUN-2003 (TREMblrel. 24, Created)
DT 01-JUN-2003 (TREMblrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE DE ForMyrmethionine deformylase (EC 3.5.1.31).
GN DEF1 OR LP 2155.
OS Lactobacillus plantarum.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus
OC Lactobacillus
OX NCBI_TaxID=1590;
[1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=NCIMB 8826 / WCFS1;
RC MEDLINE=22480296; PubMed=12566566;
RA Kleersbezem M., Boekhorst J., van Kranenburg R., Molenaar D.,
RA Kuipers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,
RA Fiers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,
RA Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing I.
RA De Vos W.M., Siezen R.J.

```

RT "Complete genome sequence of Lactobacillus plantarum WCFS1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).

DR EMBL: AL935258; CAD64511.1; -
 DR GO: GO:0008463; F:formylmethionine deformylase activity; IEA.
 DR GO: GO:0016787; F:hydrolase activity; IEA.
 DR GO: GO:0006412; P:protein biosynthesis; IEA.
 DR InterPro: IPR000181; Pep deformylase; IEA.
 DR Pfam: PF01327; Pep deformylase; 1.
 DR PRINTS: PR01576; PDEFORMLASE.
 DR Hydrolase; Complete proteome.
 SQ SEQUENCE 186 AA; 20854 MW; E080FF56D7723576 CRC64;

Query Match 49.2%; Score 488; DB 16; Length 186;
 Best Local Similarity 49.7%; Pred. No. 1.5e-35;
 Matches 91; Conservative 45; Mismatches 45; Indels 2; Gaps 1;

QY 1 MLTWKDIIRDGHPTLRQKAAAELEPLTKKEKETLIAMREFLVNSQDEETAKRYGLRSGVG 60
 Db 1 MIKWDIIRREGNPTLRVAEEVPPVITEEDRQLGDMTLTKNSQDPVKAEELQLRGVG 60
 QY 61 LAAPQINISKRMIAVLIPD--DGSKSYDYMVNPVKIVSHSVQEAAYLPTGEGCLSVDDNV 118
 Db 61 LAAPQVDVSEQAALVPSNEDEPVPKVIINPVIISHSVQCALTEGEGCLSVDRDI 120
 QY 119 AGLVHRHNKTIITKADIEGNDIQLRLKGYPAIVFQHEIDHLNGVMFYDHDIDKHDPHT 178
 Db 121 AGYVIRHDIRITLRYNNMAGEKTKRLKNYPAIVQHEIDHLHLGILFYDHLNGDNPFAADD 180
 QY 179 DAV 181
 Db 181 DLV 183

RESULT 5

ID Q82ZJ0 PRELIMINARY; PRT; 187 AA.
 AC Q82ZJ0;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Polypeptide deformylase.
 GN DEF-1 OR EF3066.
 OS Enterococcus faecalis (Streptococcus faecalis).
 OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
 OX NCBI_TaxID=1351;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX STRAIN=V587; ATCC 700802;
 RY MEDLINE=22550857; PubMed=12663927;
 RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
 RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
 RA Tettelin H., Dodson R.J., Umayam L., Brinkac L., Beanan M.,
 RA Daugherty S., DeBoy R.T., Durkin S., Kolonay J., Madupu R., Nelson W.,
 RA Vamathevan J., Tran B., Upton J., Hansen T., Shetty J., Khouri H.,
 RA Uterback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.;
 RT "Role of mobile DNA in the evolution of vancomycin-resistant
 Enterococcus faecalis.";
 RL Science 299:2071-2074(2003).
 DR EMBL: AE016956; AAC82748.1; -
 DR TIGR: EF3066; -
 DR GO: GO:0008463; F:formylmethionine deformylase activity; IEA.
 DR GO: GO:0006412; P:protein biosynthesis; IEA.
 DR InterPro: IPR000181; Pep deformylase.
 DR Pfam: PF01327; Pep deformylase; 1.
 DR PRINTS: PR01576; PDEFORMLASE.
 DR ProDom: PD003844; Pep deformylase; 1.
 DR TIGRfam: TIGR00079; pep_deformyl; 1.
 KW Complete proteome.
 SQ SEQUENCE 187 AA; 20912 MW; 9CAF46335311B0B2 CRC64;

Query Match 48.3%; Score 478.5; DB 16; Length 187;
 Best Local Similarity 55.1%; Pred. No. 1.1e-34;
 Matches 97; Conservative 30; Mismatches 46; Indels 3; Gaps 1;

QY 1 MLTWKDIIRDGHPTLRQKAAAELEPLTKKEKETLIAMREFLVNSQDEETAKRYGLRSGVG 60
 Db 1 MITMKDIIRREGNPTLRVAEEVPPVITEEDRQLGDMTLTKNSQDPVKAEELQLRGVG 60
 QY 61 LAAPQINISKRMIAVLIPD--DGSKSYDYMVNPVKIVSHSVQEAAYLPTGEGCLSVDDN 117
 Db 61 LAAPQLDISKRIIAVHVPSNDPENETSLSTVMNPKILSHSVQDVCLEGEGLSVDRD 120
 QY 118 VAGLVHRHNKTIITKADIEGNDIQLRLKGYPAIVFQHEIDHLNGVMFYDHDIDKHDP 173
 Db 121 VPGVYVHRHNKITSYFDMAGEKKHVKRLKNYPAIVQHEIDHLNGVMFYDHLNGKPNP 176

RESULT 6

ID Q842S4 PRELIMINARY; PRT; 187 AA.
 AC Q842S4;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Peptide deformylase.
 OS Enterococcus faecium (Streptococcus faecium).
 OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
 OX NCBI_TaxID=1352;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX STRAIN=ATCC6057;
 RA Bing T.X., Qin Z.Y., Yi S.S.;
 RL Submitted (FEB-2003); to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY238515; AAC88058.1; -
 DR GO: GO:0008463; F:formylmethionine deformylase activity; IEA.
 DR GO: GO:0006412; P:protein biosynthesis; IEA.
 DR InterPro: IPR000181; Pep deformylase.
 DR PRINTS: PR01576; PDEFORMLASE.
 DR ProDom: PD003844; Pep deformylase; 1.
 DR TIGRfam: TIGR00079; pep_deformyl; 1.
 SQ SEQUENCE 187 AA; 20812 MW; 49E203767AD0F257 CRC64;

Query Match 47.7%; Score 472.5; DB 2; Length 187;
 Best Local Similarity 55.7%; Pred. No. 3.6e-34;
 Matches 98; Conservative 26; Mismatches 49; Indels 3; Gaps 1;

QY 1 MLTWKDIIRDGHPTLRQKAAAELEPLTKKEKETLIAMREFLVNSQDEETAKRYGLRSGVG 60
 Db 1 MITMKDIIRREGNPTLRVAEEVPPVITEEDRQLGDMTLTKNSQDPVKAEELHLRGVG 60
 QY 61 LAAPQINISKRMIAVLIPD--DGSKSYDYMVNPVKIVSHSVQEAAYLPTGEGCLSVDDN 117
 Db 61 LAAPQLDISKRIIAVHVPSNDPENETSLSTVMNPKILSHSVQDVCLEGEGLSVDRD 120
 QY 118 VAGLVHRHNKTIITKADIEGNDIQLRLKGYPAIVFQHEIDHLNGVMFYDHDIDKHDP 173
 Db 121 VPGVYVHRHNKITSYFDMAGEKKHVKRLKNYPAIVQHEIDHLNGVMFYDHLNGKPNP 176

RESULT 7

ID Q8E378 PRELIMINARY; PRT; 204 AA.
 AC Q8E378;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein.
 GN GSI1883.
 OS Streptococcus agalactiae (serotype III).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OX NCBI_TaxID=216495;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX STRAIN=NEM316 / Serotype III;

RX MEDLINE=22242508; PubMed=12354221;
 RA Glaser P., Rusnick C., Buchrieser C., Chevalier F., Frangeul L.,
 RA Meadek T., Zouine M., Couve E., Lalioui L., Poyart C., Irieu-Cuot P.,
 RA Kunst F.;
 RT "Genome sequence of *Streptococcus agalactiae*, a pathogen causing
 RT invasive neonatal disease."
 RL Mol. Microbiol. 45:1499-1513 (2002).
 DR EMBL; AL766854; CAD47542.1; -;
 DR SAGEList; gbal883; -;
 DR GO; GO:0008463; F:formylmethionine deformylase activity; IEA.
 DR GO; GO:0006442; P:protein biosynthesis; IEA.
 DR InterPro; IPR000181; Pep deformylase.
 DR Pfam; PF01327; P:deformylase.
 DR PRINTS; PR01576; P:DEFORMYLASE.
 DR ProDom; PD003844; Pep deformylase; 1.
 DR TIGRFAMS; TIGR00079; pep_deformyl; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 204 AA; 22830 MW; 50097F6CCF8524EF CRC64;

 Query Match 45.7%; Score 453; DB 16; Length 204;
 Best Local Similarity 51.4%; Pred. No. 2.2e-32;
 Matches 93; Conservative 36; Mismatches 44; Indels 8; Gaps 2;

 QY 1 MLTKDIIIRDGHPTLRKAAELEPLTKKEKTLIAMREFLVNSQDEIARVGLRSGVG 60
 DB 13 LIDNDIIRGNPTLRKVAEVTPLSEKEILGEKMMQFLKHSQDPIAEKGLRSGVG 72
 QY 61 LAAPQINISKRMIAVLIP--DDGSGK-----SYDYMVNPKIVSHSVQEAAYLTGSGCL 112
 DB 73 LAAPQLDISKRIIAVLVPNVEDAQNPPKEAYSLSQEVWYNPKVSHSVQDAALSDGSGCL 132
 QY 113 SVDNVAGLVHRHNKITIKAKDIEGNDIQLRLKGYPAIVFOHEIDHLNGVMFYDHIDKH 172
 DB 133 SVDREPGYVVRHARVTIEYFDKTEGKHLKLGKGYNSIVVQHEIDHIDGIMFYDRINEKN 192
 QY 173 P 173
 DB 193 P 193

 RESULT 8
 Q8DXF6 PRELIMINARY; PRT; 204 AA.
 AC Q8DXF6;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DE Polypeptide deformylase.
 GN DEF OR SAG1895.
 OS *Streptococcus agalactiae* (serotype V).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=216466;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2603 V/R / Serotype V;
 RX MEDLINE=22222986; PubMed=12200547;
 RA Tettelin H., Massignani V., Cieslewicz M.J., Eisen J.A., Peterson S.,
 RA Wessels M.R., Paulsen I.T., Nelson K.E., Margat I., Read T.D.,
 RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
 RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
 RA Radune D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,
 RA Carthy H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
 RA Iacobino E.T., Brettoni C., Galli G., Mariani M., Vegni F., Malone D.,
 RA Rinaudo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
 RA Fraser C.M.;
 RT "Complete genome sequence and comparative genomic analysis of an
 RT emerging human pathogen, serotype V *Streptococcus agalactiae*."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396 (2002).
 DR EMBL; AE014277; AAN00757.1; -;
 DR TIGR; SAG1895; -;
 DR GO; GO:0008463; F:formylmethionine deformylase activity; IEA.
 DR GO; GO:0006442; P:protein biosynthesis; IEA.

DR InterPro; IPR000181; Pep deformylase.
 DR Pfam; PF01327; Pep deformylase; 1.
 DR PRINTS; PR01576; P:DEFORMYLASE.
 DR ProDom; PD003844; Pep deformylase; 1.
 DR TIGRFAMS; TIGR00079; pep_deformyl; 1.
 KW Complete proteome.
 SQ SEQUENCE 204 AA; 22830 MW; 50097F6CCF8524EF CRC64;

 Query Match 45.7%; Score 453; DB 16; Length 204;
 Best Local Similarity 51.4%; Pred. No. 2.2e-32;
 Matches 93; Conservative 36; Mismatches 44; Indels 8; Gaps 2;

 QY 1 MLTKDIIIRDGHPTLRKAAELEPLTKKEKTLIAMREFLVNSQDEIARVGLRSGVG 60
 DB 13 LIDNDIIRGNPTLRKVAEVTPLSEKEILGEKMMQFLKHSQDPIAEKGLRSGVG 72
 QY 61 LAAPQINISKRMIAVLIP--DDGSGK-----SYDYMVNPKIVSHSVQEAAYLTGSGCL 112
 DB 73 LAAPQLDISKRIIAVLVPNVEDAQNPPKEAYSLSQEVWYNPKVSHSVQDAALSDGSGCL 132
 QY 113 SVDNVAGLVHRHNKITIKAKDIEGNDIQLRLKGYPAIVFOHEIDHLNGVMFYDHIDKH 172
 DB 133 SVDREPGYVVRHARVTIEYFDKTEGKHLKLGKGYNSIVVQHEIDHIDGIMFYDRINEKN 192
 QY 173 P 173
 DB 193 P 193

 RESULT 9
 Q939R9 PRELIMINARY; PRT; 203 AA.
 AC Q939R9;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DE Peptide deformylase DefB (EC 3.5.1.88) (PDF) (Polypeptide
 DE deformylase).
 GN DEF OR DEF8.
 OS *Streptococcus pneumoniae*.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=R6x;
 RX MEDLINE=21393646; PubMed=11502510;
 RA Margolis P., Hackbarth C., Lopez S., Maniar M., Wang W., Yuan Z.,
 RA White R., Trias J.;
 RT "Resistance of *Streptococcus pneumoniae* to deformylase inhibitors is
 RT due to mutations in defB".
 RL Antimicrob. Agents Chemother. 45:2432-2435 (2001).
 CC -I- FUNCTION: REMOVES THE FORMYL GROUP FROM THE N-TERMINAL MET OF
 CC NEWLY SYNTHESIZED PROTEINS. REQUIRES AT LEAST A DIPEPTIDE FOR AN
 CC EFFICIENT RATE OF REACTION. N-TERMINAL L-METHIONINE IS A
 CC PREREQUISITE FOR ACTIVITY BUT THE ENZYME HAS BROAD SPECIFICITY AT
 CC OTHER POSITIONS (BY SIMILARITY).
 CC -I- CATALYTIC ACTIVITY: FORMYL-L-METHIONYL PEPTIDE + H(2)O = FORMATE +
 CC METHIONYL PEPTIDE.
 CC -I- COFACTOR: BINDS 1 IRON(II) ION (BY SIMILARITY).
 CC -I- SIMILARITY: BELONGS TO THE POLYPEPTIDE DEFORMYLASE FAMILY.
 DR EMBL; AY014509; AAK1328.1; -;
 DR PIR; E98035; E98035.
 DR GO; GO:0008463; F:formylmethionine deformylase activity; IEA.
 DR PIR; GO:0016787; F:hydrolase activity; IEA.
 DR GO; GO:0042586; F:peptide deformylase activity; IEA.
 DR GO; GO:0006412; P:protein biosynthesis; IEA.
 DR InterPro; IPR000181; Pep deformylase.
 DR Pfam; PF01327; Pep deformylase; 1.
 DR PRINTS; PR01576; P:DEFORMYLASE.
 DR ProDom; PD003844; Pep deformylase; 1.
 DR TIGRFAMS; TIGR00079; pep_deformyl; 1.
 DR Hydrolase; Iron; Protein biosynthesis.

SEQ	SEQUENCE	203 AA; 22692 MW; E332956982A67161 CRC64;	45.6%; Score 451.5; DB 2; Length 203; Best Local Similarity 51.7%; Pred. No. 3e-32; Matches 93; Conservative 35; Mismatches 45; Indels 7; Gaps 2;
QY	1	MLTMKDIIIRDGHPTLRQKAABLEPLTKEEKETLIAMREFLVNSQDEEIIAKRYGLRSGVG	60
Db	13	LIDMNDIIRREGNPTLRVAVBEVTFPLSDQETILGKMMQFLKHSQDPVMAEKMLRGSGVG	72
QY	61	LAAPQINISKRMIAVLIPD-----DSGGKSYDY--MLVNPKIYSHSVQEAYLPTGEGCLS	113
Db	73	LAAPQLDISKRIIAVLVFNIVEGETPQEAIDYDAIMYNPKIYSHSVQDAALGEGEGCLS	132
QY	114	VDDNVAGLVHRHNKTIKAKDIEGNDQLRLKGYPALVFQHEIDHLNGVMFYDHDIDKDP	173
Db	133	VDRNVPGVYVHRVTVDYEDKQGEKIRIKLKGYSINIVQHEIDHLNGIMFYDRINEKDP	192
RESULT 10			
Q8DP79		PRELIMINARY; PRT; 203 AA.	
ID	Q8DP79		
AC	Q8DP79;		
DT	01-MAR-2003 (TrEMBLrel. 23, Created)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)		
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)		
DE	Peptide deformylase, N-formylmethionylaminoacyl-tRNA deformylase (EC		
DE	3.5.1.31)		
GN	FMS OR SPRI1310.		
OS	Streptococcus pneumoniae (strain ATCC BAA-255 / R6).		
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;		
OC	Streptococcus.		
OX	NCBI_TaxID=171101;		
ON	[1]		
RP	SEQUENCE FROM N.A.		
RP	MEDLINE=21429245; PubMed=11544234;		
RA	Hoskins J., Alboin W.B. Jr., Arnold J., Blaszcak L.C., Burgett S.,		
RA	DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,		
RA	Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,		
RA	LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,		
RA	McAhren S.M., McHenney M., McLeaster K., Mundy C.W., Nicas T.I.,		
RA	Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rokey P.,		
RA	Sun F.-M., Winkler M.B., Yang Y., Young-Bellido M., Zhao G.,		
RA	Zook C.A., Baltz R.H., Jaskunas S.R., Rosteck P.R. Jr., Skatrud P.L.,		
RA	Glaes J.I.;		
RT	"Genome of the bacterium Streptococcus pneumoniae strain R6.";		
RL	J. Bacteriol. 183:5709-5717(2001).		
RL	EMBL; A8008502; AAL00114.1; -.		
DR	PIR; E98035; E98035.		
DR	GO; GO:0008463; F:formylmethionine deformylase activity; IEA.		
DR	GO; GO:0016787; F:hydrolase activity; IEA.		
DR	GO; GO:0006412; P:protein biosynthesis; IEA.		
DR	InterPro; IPR000181; Pep deformylase.		
DR	Pfam; PF01327; Pep deformylase; 1.		
DR	PRINTS; PR01576; PDEFORMLASE.		
DR	ProDom; PD003844; Pep deformylase; 1.		
DR	TIGRFAMS; TIGR00079; pept deformyl; 1.		
KW	Hydrolase; Complete proteome.		
SEQ	SEQUENCE 203 AA; 22692 MW; E332956982A67161 CRC64;		
Query Match		45.6%; Score 451.5; DB 16; Length 203;	
Best Local Similarity		51.7%; Pred. No. 3e-32;	
Matches		93; Conservative 35; Mismatches 45; Indels 7; Gaps 2;	
QY	1	MLTMKDIIIRDGHPTLRQKAABLEPLTKEEKETLIAMREFLVNSQDEEIIAKRYGLRSGVG	60
Db	13	LIDMNDIIRREGNPTLRVAVBEVTFPLSDQETILGKMMQFLKHSQDPVMAEKMLRGSGVG	72
QY	61	LAAPQINISKRMIAVLIPD-----DSGGKSYDY--MLVNPKIYSHSVQEAYLPTGEGCLS	113
Db	73	LAAPQLDISKRIIAVLVFNIVEGETPQEAIDYDAIMYNPKIYSHSVQDAALGEGEGCLS	132
QY	114	VDDNVAGLVHRHNKTIKAKDIEGNDQLRLKGYPALVFQHEIDHLNGVMFYDHDIDKDP	173


```

RESULT 15
Q8LEHO PRELIMINARY; PRT; 273 AA.
AC Q8LEHO;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Formylmethionine deformylase.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eutrosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation.";
RL Genome Biol. 0:0-0(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AY085417; AAM62644.1; -.
DR GO; GO:0008463; F:formylmethionine deformylase activity; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR InterPro; IPR000181; Pep_deformylase.
DR Pfam; PF01327; Pep_deformylase; 1.
DR PRINTS; PR01576; PDEFORMYLASE.
DR ProDom; PD003844; Pep_deformylase; 1.
DR TIGRfams; TIGR00079; pep_deformyl; 1.
SQ SEQUENCE 273 AA; 30623 MW; 754ACDD231B436EB CRC64;

Query Match 20.5%; Score 203.5; DB 10; Length 273;
Best Local Similarity 34.4%; Pred. No. 5.7e-10;
Matches 52; Conservative 27; Mismatches 67; Indels 5; Gaps 4;

QY 18 KAAELEPLTKKEETLIAMREFLVNSQDEEIAKRYGLRSGVGLAAPQINISKRMIAVLI 77
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
82 KIVEYPPILRAKMKRIDIFDENLKNLVDAWDFVWY-KTDGIGLSAPQVGLNQLM-VFN 139
QY 78 PDGSGSKSYDMLVNPXIVSHSVCEAYLPTGEGCLSVDDNVAGLVHRHNKITIKAKDIEG 137
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
140 PAGEPGEKKEIVLVNPKIKKYS--DKLVPFDEGCLSPFGIYAEVV-RPQSVKIDARDITG 196
QY 138 NDIQLRLKGYPAIVFQHEIDHLNGVMFYDHI 168
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
197 ERFSSISLSRLPARIFQHEYDHLGVLFFDRM 227

```

Search completed: March 31, 2004, 16:36:39
Job time : 55 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 16:34:10 ; Search time 23 Seconds
(without alignments)
424.231 Million cell updates/sec

Title: US-09-896-580B-12
Perfect score: 991
Sequence: 1 MLTWKDIIRGHPTLRQAA.....KDHPLQPHDTAVEVHQHHH 189

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA.*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pap.*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pap.*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pap.*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pap.*
5: /cgn2_6/ptodata/2/iaa/PCTUS.COMB.pap.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	938	94.7	183	4	US-09-373-953-2
2	938	94.7	183	4	US-08-911-844B-2
3	795	80.2	210	4	US-09-134-001C-3446
4	754	76.1	150	4	US-09-373-953-4
5	754	76.1	150	4	US-08-911-844B-4
6	446.5	45.1	203	3	US-08-991-023-2
7	386.5	39.0	155	3	US-08-991-023-4
8	214	21.6	176	4	US-03-543-681A-6922
9	209.5	21.1	176	4	US-03-107-532A-6622
10	198.5	20.0	164	4	US-09-328-352-5854
11	198	20.0	181	2	US-08-932-142-2
12	198	20.0	181	4	US-09-342-458-2
13	186.5	18.8	160	4	US-03-194-146-8
14	180.5	18.2	169	2	US-08-955-339-4
15	180.5	18.2	169	3	US-09-188-820-4
16	177	17.9	175	4	US-09-489-039A-10978
17	176	17.8	203	4	US-09-328-352-4826
18	175	17.7	192	4	US-09-252-091A-21776
19	170.5	17.2	172	4	US-03-543-681A-7238
20	169.5	17.1	204	4	US-03-198-452A-4
21	155	15.6	194	4	US-09-540-236-2156
22	154.5	15.6	185	4	US-09-252-991A-19647
23	147	14.8	163	4	US-09-134-001C-4828
24	129.5	13.1	151	4	US-09-134-000C-4110
25	88	8.9	455	4	US-09-540-236-2325
26	87.5	8.8	290	4	US-09-328-352-7390
27	83.5	8.4	555	4	US-09-252-991A-30276

28 81 8.2 638 2 US-08-557-122A-38 Sequence 38, Appl
29 81 8.2 638 4 US-09-262-666-38 Sequence 38, Appl
30 81 8.2 769 4 US-09-252-991A-17737 Sequence 17737, A
31 81 8.2 1447 4 US-09-376-330-17 Sequence 17, Appl
32 80.5 8.1 612 4 US-09-489-039A-10140 Sequence 10140, A
33 80.5 8.1 896 4 US-09-689-085B-5 Sequence 5, Appl
34 79 8.0 563 4 US-09-540-236-2399 Sequence 2939, Ap
35 79 8.0 724 4 US-09-328-352-7710 Sequence 7710, Ap
36 78.5 7.9 604 1 US-08-487-753-2 Sequence 2, Appl
37 78.5 7.9 604 2 US-08-480-065-2 Sequence 2, Appl
38 78.5 7.9 604 3 US-08-487-744-2 Sequence 2, Appl
39 78.5 7.9 604 5 PCT-US93-09167-2 Sequence 2, Appl
40 77 7.8 1061 4 US-09-762-481B-2 Sequence 2, Appl
41 76.5 7.7 523 4 US-09-604-957-5 Sequence 5, Appl
42 76.5 7.7 715 4 US-09-543-681A-5291 Sequence 5291, Ap
43 76 7.7 580 4 US-09-328-352-7656 Sequence 7656, Ap
44 76 7.7 928 4 US-09-914-259-23 Sequence 23, Appl
45 76 7.7 1079 4 US-09-489-039A-7502 Sequence 7502, Ap

ALIGNMENTS

RESULT 1

US-09-373-953-2
; Sequence 2, Application US/09373953
; Patent No. 6410868

; GENERAL INFORMATION:
; APPLICANT: Lonetto, Michael A
; APPLICANT: Sylvester, Daniel
; APPLICANT: Warren, Richard

; TITLE OF INVENTION: No. 6410688e1 defL
; FILE REFERENCE: GM10001

; CURRENT APPLICATION NUMBER: US/09/373,953
; CURRENT FILING DATE: 1999-08-13

; PRIOR APPLICATION NUMBER: US 08/911,844
; PRIOR FILING DATE: 1997-08-15

; PRIOR APPLICATION NUMBER: US 60/048,706
; PRIOR FILING DATE: 1997-05-21

; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2
; LENGTH: 183

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-09-373-953-2

Query Match 94.7%; Score 938; DB 4; Length 183;
Best Local Similarity 98.9%; Pred. No. 1.5e-97;
Matches 181; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLTWKDIIRGHPTLRQAAAELEPLTKKEKFTLIAMREFLVNSQDEIARYGLRSGVG 60

Db 1 MLTWKDIIRGHPTLRQAAAELEPLTKKEKFTLIAMREFLVNSQDEIARYGLRSGVG 60

QY 61 LAAPQINISKRMIAVLIPDDGSGKSDYMLVNPKI VSHSVQEA VLP TGECLSVDDNVAG 120

Db 61 LAAPQINISKRMIAVLIPDDGSGKSDYMLVNPKI VSHSVQEA VLP TGECLSVDDNVAG 120

QY 121 LVHRHNKTTIKAKDIEGNDIQLKGYPAIVFOHEIDHLNGVMFYDHIKDXHPLQPHHTDA 180

Db 121 LVHRHNKTTIKAKDIEGNDIQLKGYPAIVFOHEIDHLNGVMFYDHIKDXHPLQPHHTDA 180

QY 181 VEV 183

Db 181 VEV 183

RESULT 2

US-08-911-844B-2

; Sequence 2, Application US/08911844B

; Patent No. 6586578

; GENERAL INFORMATION:

```

; APPLICANT: Lonetto, Michael A
; APPLICANT: Sylvester, Daniel
; APPLICANT: Warren, Richard
; TITLE OF INVENTION: No. 6586578el defL
; FILE REFERENCE: GM10001
; CURRENT APPLICATION NUMBER: US/08/911,844B
; PRIOR FILING DATE: 1997-08-15
; PRIOR APPLICATION NUMBER: US 60/048,706
; PRIOR FILING DATE: 1997-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-08-911-844B-2

Query Match      94.7%; Score 938; DB 4; Length 183;
Best Local Similarity 98.9%; Pred. No. 1.5e-97;
Matches 181; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1  MLTMKDIIRDGHPTLRQKAAELPLTKKEKETLIAMREFLVNSQDEETAKRYGLRSGVG 60
Db 1  MLTMKDIIRDGHPTLRQKAAELPLTKKEKETLIAMREFLVNSQDEETAKRYGLRSGVG 60
QY 61  LAAPQINISKRMIAVLIPDDGSGKSYDYMLNPKIVSHSVQEAYLPTGEGCLSVDDNVAG 120
Db 61  LAAPQINISKRMIAVLIPDDGSGKSYDYMLNPKIVSHSVQEAYLPTGEGCLSVDDNVAG 120
QY 121  LVHRNKITIKAKDIEGNDIQRLKGYPAIVQHEIDLHNGVMFYDIDKHPLQPHDTA 180
Db 121  LVHRNKITIKAKDIEGNDIQRLKGYPAIVQHEIDLHNGVMFYDIDKHPLQPHDTA 180
QY 181  VEV 183
Db 181  VEV 183

RESULT 3
US-09-134-001C-3446
; Sequence 3446, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3446
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3446

Query Match      80.2%; Score 795; DB 4; Length 210;
Best Local Similarity 79.8%; Pred. No. 2.3e-81;
Matches 146; Conservative 23; Mismatches 14; Indels 0; Gaps 0;

QY 1  MLTMKDIIRDGHPTLRQKAAELPLTKKEKETLIAMREFLVNSQDEETAKRYGLRSGVG 60
Db 28  MITMKDIIRDGHPTLRQKAAELPLTKKEKETLIAMREFLVNSQDEETAKRYGLRSGVG 87
QY 61  LAAPQINISKRMIAVLIPDDGSGKSYDYMLNPKIVSHSVQEAYLPTGEGCLSVDDNVAG 120
Db 88  LAAPQINISKRMIAVLIPDDGSGKSYDYMLNPKIVSHSVQEAYLPTGEGCLSVDDNVAG 147
QY 121  LVHRNKITIKAKDIEGNDIQRLKGYPAIVQHEIDLHNGVMFYDIDKHPLQPHDTA 180
```

```

; APPLICANT: Lonetto, Michael A
; APPLICANT: Sylvester, Daniel
; APPLICANT: Warren, Richard
; TITLE OF INVENTION: No. 6410688el defL
; FILE REFERENCE: GM10001
; CURRENT APPLICATION NUMBER: US/09/373,953
; CURRENT FILING DATE: 1999-08-13
; PRIOR APPLICATION NUMBER: US 08/911,844
; PRIOR FILING DATE: 1997-08-15
; PRIOR APPLICATION NUMBER: US 60/048,706
; PRIOR FILING DATE: 1997-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-373-953-4

Query Match      76.1%; Score 754; DB 4; Length 150;
Best Local Similarity 98.7%; Pred. No. 5.6e-77;
Matches 148; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1  MLTMKDIIRDGHPTLRQKAAELPLTKKEKETLIAMREFLVNSQDEETAKRYGLRSGVG 60
Db 1  MLTMKDIIRDGHPTLRQKAAELPLTKKEKETLIAMREFLVNSQDEETAKRYGLRSGVG 60
QY 61  LAAPQINISKRMIAVLIPDDGSGKSYDYMLNPKIVSHSVQEAYLPTGEGCLSVDDNVAG 120
Db 61  LAAPQINISKRMIAVLIPDDGSGKSYDYMLNPKIVSHSVQEAYLPTGEGCLSVDDNVAG 120
QY 121  LVHRNKITIKAKDIEGNDIQRLKGYPAI 150
Db 121  LVHRNKITIKAKDIEGNDIQRLKGYPAI 150

RESULT 5
US-08-911-844B-4
; Sequence 4, Application US/08911844B
; Patent No. 6586578
; GENERAL INFORMATION:
; APPLICANT: Lonetto, Michael A
; APPLICANT: Sylvester, Daniel
; APPLICANT: Warren, Richard
; TITLE OF INVENTION: No. 6586578el defL
; FILE REFERENCE: GM10001
; CURRENT APPLICATION NUMBER: US/08/911,844B
; CURRENT FILING DATE: 1997-08-15
; PRIOR APPLICATION NUMBER: US 60/048,706
; PRIOR FILING DATE: 1997-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-08-911-844B-4

Query Match      76.1%; Score 754; DB 4; Length 150;
Best Local Similarity 98.7%; Pred. No. 5.6e-77;
```

Matches 148; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLTMDIIRDGHTLRQKAABELEPLTKKEKTLTAMREFLVNSQDEETAKYGLRSGVG 60
 Db 1 MLTMDIIRDGHTLRQKAABELEPLTKKEKTLTAMREFLVNSQDEETAKYGLRSGVG 60
 QY 61 LAAPQINISKRMIAVLIPDGGSGKSYDYMLVNPVKIVSHSVQAYLPTGCGCLSVDDNVAG 120
 Db 61 LAAPQINISKRMIAVLIPDGGSGKSYDYMLVNPVKIVSHSVQAYLPTGCGCLSVDDNVAG 120
 QY 121 LVHRNKITIKAKDIEGNDIQLRLKGYPAI 150
 Db 121 LVHRNRITIKAKDIEGNDIQLRLKGYPAI 150

RESULT 6

US-08-991-023-2
 ; Sequence 2, Application US/08991023
 ; Patent No. 6284878
 ; GENERAL INFORMATION:
 ; APPLICANT: Lonetto, Michael A.
 ; TITLE OF INVENTION: def1
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dechert Price & Rhoads
 ; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
 ; CITY: Philadelphia
 ; STATE: PA
 ; COUNTRY: US
 ; ZIP: 19103
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/991,023
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Falk, Stephen T
 ; REGISTRATION NUMBER: 36,795
 ; REFERENCE/DOCKET NUMBER: GM50010
 ; TELEPHONE: 215-994-2488
 ; TELEFAX: 215-994-2222
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 203 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear

US-08-991-023-2

Query Match 45.1%; Score 446.5; DB 3; Length 203;
 Best Local Similarity 51.1%; Pred. No. 3.5e-42;
 Matches 92; Conservative 35; Mismatches 46; Indels 7; Gaps 2;
 QY 1 MLTMDIIRDGHTLRQKAABELEPLTKKEKTLTAMREFLVNSQDEETAKYGLRSGVG 60
 Db 13 LIDMNDIIRGNPXLRTVAEVTFLSDQBIILGKMMQFLKHSQDPVMAEKMLRGVG 72
 QY 61 LAAPQINISKRMIAVLIPDGGSGKSYDYMLVNPVKIVSHSVQAYLPTGCGCLSV 113
 Db 73 LAAPQIDISKRIITAVLPNIIVEGETPQEAAYDLEAIMNPVKIVSHSVQDAALGEGGCL 132
 QY 114 VDNVAGLVHRNKITIKAKDIEGNDIQLRLKGYPAI VFOHEIDHLNGVMFYDHIDKHP 173
 Db 133 VDRNVGYVVRHARVTVDFDKGKHKRIKLGKYNLSVQVQHEIDHLNGVMFYDHIDKHP 192

RESULT 7

US-08-991-023-4
 ; Sequence 4, Application US/08991023
 ; Patent No. 6284878
 ; GENERAL INFORMATION:
 ; APPLICANT: Lonetto, Michael A.
 ; TITLE OF INVENTION: def1
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dechert Price & Rhoads
 ; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
 ; CITY: Philadelphia
 ; STATE: PA
 ; COUNTRY: US
 ; ZIP: 19103
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/991,023
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Falk, Stephen T
 ; REGISTRATION NUMBER: 36,795
 ; REFERENCE/DOCKET NUMBER: GM50010
 ; TELEPHONE: 215-994-2488
 ; TELEFAX: 215-994-2222
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 155 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear

US-08-991-023-4
 Query Match 39.0%; Score 386.5; DB 3; Length 155;
 Best Local Similarity 54.2%; Pred. No. 1.3e-35;
 Matches 78; Conservative 28; Mismatches 31; Indels 7; Gaps 2;
 QY 37 MREFLVNSQDEETAKYGLRSGVGLAAPQINISKRMIAVLIPD-----DGGKSYDY--M 89
 Db 1 MMOLFHSQDPVMAEKMLRGVGLAAPQIDISKRIITAVLPNIIVEGETPQEAAYDLEAI 60
 QY 90 LVNPVKIVSHSVQAYLPTGCGCLSVDDNVAGLVHRNKITIKAKDIEGNDIQLRLKGYPA 149
 Db 61 MYNPVKIVSHSVQDAALGEGGCLSVDRNPGYVVRHARVTVDFDKGKHKRIKLGKYN 120
 QY 150 VFOHEIDHLNGVMFYDHIDKHP 173
 Db 121 VVQHEIDHLNGVMFYDHIDKHP 144

RESULT 8

US-09-543-681A-6922
 ; Sequence 6922, Application US/09543681A
 ; Patent No. 6605709
 ; GENERAL INFORMATION:
 ; APPLICANT: GARY BRETON
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABIL
 ; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.1002-001
 ; CURRENT APPLICATION NUMBER: US/09/543,681A
 ; CURRENT FILING DATE: 2000-04-05

;; PRIOR APPLICATION NUMBER: US 60/128,706
;; PRIOR FILING DATE: 1999-04-09
;; NUMBER OF SEQ ID NOS: 8344
;; SEQ ID NO 6922
;; LENGTH: 176
;; TYPE: PRT
;; ORGANISM: Proteus mirabilis
US-09-543-681A-6922

Query Match 21.6%; Score 214; DB 4; Length 176;
Best Local Similarity 34.8%; Pred. No. 4.1e-16;
Matches 56; Conservative 29; Mismatches 54; Indels 22; Gaps 7;
QY 15 LRQKAAELPLTKKEETLIAMREFLVNSQDEIAKRYGLRSVGLAAPQINISKRMIA 74
Db 18 LRKVAPE--KVDDIRTL-----DDMIETMAER-GIGLAAPQNVSKRIYV 64
QY 75 VLIPDGSKSYDMLVNPVKIVSHSVQEAFLPTGEGCLSVDDNVAGLVHRHNKTIKAKD 134
Db 65 I--DVSENRDOPALINPEII--STEDEIMDMMDGCLSIPIPSFAP-TQFRYLYKVKALD 118
QY 135 IEGNDIQRLKGYPAIVFOHEIDHLNGVMFYDHIKDKHPLQ 175
Db 119 RNGDEIELEADLEFAGCIQHEDHLNGKLFIDHLS---ELK 156

RESULT 9
US-09-107-532A-6622
; Sequence 6622, Application US/09107532A
; Patent No. 6563275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Denise
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6622:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 176 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature

;; LOCATION: (B) LOCATION 1...176
;; SEQUENCE DESCRIPTION: SEQ ID NO: 6622:
US-09-107-532A-6622

Query Match 21.1%; Score 209.5; DB 4; Length 176;
Best Local Similarity 30.9%; Pred. No. 1.3e-15;
Matches 59; Conservative 39; Mismatches 64; Indels 29; Gaps 7;
QY 3 TMKDIIRDG-----HPTLRQKAAELEPLTKKEETLIAMREFLVNSQDEETAKRY 53
Db 3 TIKLYLRKQIMRYPIILHPNDKLRKTAQPIDVITDETIAL-----LDNLVETMIA--- 53
QY 54 GLRSVGLAAPQINISKRMIAVLIPDDSGSKSYDMLVNPVKIVSHSVQEAFLPTGEGCLS 113
Db 54 --NDGIGIAAPQVGNKR-IAVIEVDEGE---KPELINPEIEIAKGSLOV---EGCLS 103
QY 114 VDDNVAGLVHRHNKTIKAKDIEGNDIQRLKGYPAIVFOHEIDHLNGVMFYDHIKDKHP 173
Db 104 I-PHYVGVTKRADEVTVRYDREDGEEIEVTAAGYLARAFQHEIDHLDGILFIKMIQIIP 162
QY 174 LQHTDAVEVH 184
Db 163 EEELEEYMEEH 173

RESULT 10
US-09-328-352-5854
; Sequence 5854, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03FA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5854
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5854

Query Match 20.0%; Score 198.5; DB 4; Length 164;
Best Local Similarity 38.7%; Pred. No. 2.1e-14;
Matches 46; Conservative 23; Mismatches 39; Indels 11; Gaps 4;
QY 56 RSGVGLAAPQINISKRMIAVL-----IPDDSGSKSYDMLVNPVKIVSHSVQEAFLPTGE 109
Db 50 RNVGVIAAPQVYISKRVIIIVASRPNRYPD--APENNAVVMVNPFIIEFSSETCL--GEE 105
QY 110 GCLSVDDNVAGLVHRHNKTIKAKDIEGNDIQRLKGYPAIVFOHEIDHLNGVMFYDHI 168
Db 106 GCLSVDPDE-RQQVERAEMVKYLTLOGEAVETIFHGFPARIVQHEVDHLNGILFVERI 163

RESULT 11
US-08-932-142-2
; Sequence 2, Application US/08932142
; Patent No. 5962666
; GENERAL INFORMATION:
; APPLICANT: Lonetto, Michael A.
; TITLE OF INVENTION: NOVEL def
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette

```

;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2252
; TELEFAX: 215-994-2222
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 181 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-342-458-2

Query Match          20.0%; Score 198; DB 4; Length 181;
Best Local Similarity 33.5%; Pred. No. 2.7e-14;
Matches 60; Conservative 31; Mismatches 62; Indels 26; Gaps 8;

QY      7 IIRD---GHPTLRQKAAAELELPITKKEKETLIAMREFLVNSODEETAKYGLRSGVGLA 62
      :|||:      :|||:      :|||:      :|||:      :|||:      :|||:
Db      1 MIRDLEYDSPILRKVAAPV-TEITDELQRLQVLDMSETM-----AFYKGVGLA 47

QY      63 APOINISKMIIVLIP---DDGSGKSYDY--MLVNPKIVSHSVQEAYLPTGEGCLSVDDN 117
      :|||:      :|||:      :|||:      :|||:      :|||:      :|||:
Db      48 APOVGOSISLIFNGVERELEDGELVFCDFPRVFINPVITQKSEQLVY--GNEGCLSI-LR 104

QY      118 VAGLVHRHNKITTAKADIEGNDIQLRLKGYPAIVFQHEIDLHNGVMFYDHI-DKHPLQ 175
      :|||:      :|||:      :|||:      :|||:      :|||:      :|||:
Db      105 IRGEVARPDKITESAKNLDGQQLSLAEGFLARIVMHEIDHLHGVLIDRMSDKDKTKQ 163

RESULT 13
US-09-194-146-8
; Sequence 8, Application US/09194146
; Patent No. 6458557
; GENERAL INFORMATION:
; APPLICANT: Miller, Brian
; APPLICANT: Diaz-Torres, Maria
; TITLE OF INVENTION: Gram-Positive Microorganism Formate Pathway
; FILE REFERENCE: GC395-US
; CURRENT APPLICATION NUMBER: US/09/194,146
; CURRENT FILING DATE: 1998-11-19
; PRIOR APPLICATION NUMBER: GB 9724627.6
; PRIOR FILING DATE: 1997-11-20
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Bacillus
US-09-194-146-8

Query Match          18.8%; Score 186.5; DB 4; Length 160;
Best Local Similarity 35.7%; Pred. No. 4.5e-13;
Matches 61; Conservative 26; Mismatches 59; Indels 25; Gaps 8;

QY      2 LTMKDIIRDGHT-IRQKAAEELPITKKEKETLIAMREFLVNSODEETAKYGLRSGV 60
      :|||:      :|||:      :|||:      :|||:      :|||:      :|||:
Db      1 LAVKKVVT--HPAEVLETPAETVTVFDKKLKLDDMYDTMLE-----MDGVG 46

QY      61 LAAPQINISKMIIVLIPDDGSGKSYDYMLVNPKIVSHSVQEAYLPTG-EGCLSVDDNVA 119
      :|||:      :|||:      :|||:      :|||:      :|||:      :|||:
Db      47 LAAPQIGLKRAAVVEIGD-RGR---IDLNPKELEKSGEQ-----TGIEGCLSF-PNVY 97

QY      120 GLVHRHNKITTAKADIEGNDIQLRLKGYPAIVFQHEIDLHNGVMFYDHIK 170
      :|||:      :|||:      :|||:      :|||:      :|||:      :|||:
Db      98 GDVTRADYVKVRAFNRQCKPFLEARGFLARAVQHEMDHLDGVLFTSKISK 148

RESULT 14
US-08-895-939-4
; Sequence 4, Application US/08895939
; Patent No. 5834243
; GENERAL INFORMATION:

```

```

; APPLICANT: Bogosian, Gregg
; TITLE OF INVENTION: Deforylation of f-Met Peptides in
; Bacterial Expression Systems
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.
; ZIP: 77210
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/08/895,939
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/022,687
; FILING DATE: 17-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Patterson, Melinda L.
; REGISTRATION NUMBER: 33,062
; REFERENCE/DOCKET NUMBER: MOPV:006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 787-1400
; TELEFAX: (713) 789-2679
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 169 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-895-939-4

Query Match 18.2%; Score 180.5; DB 2; Length 169;
Best Local Similarity 32.9%; Pred. No. 2.3e-12;
Matches 53; Conservative 31; Mismatches 54; Indels 23; Gaps 8;

Qy 22 LELPLTKKEKETLIAMREFLVNSQ-----DEEIAKRYGLRSGVGLAAPQINISKMTIAVL 76
Db 7 LHIP---DERLRKVKAPVEEVNAEIQRIVDDMFETMYA-BEGIGLAATQVDIHQRIIVI- 61
Qy 77 IPDDGSGKSYDYMVLNPKIVSHSVQEAYLPTG--EGCLSVDDNVAGLVHRHNKITIKAKD 134
Db 62 --DVSENRDERLVLINPELLEKSGE-----TGIEEGCLSIPEQRA-LVPRAEKVKIRALD 113
Qy 135 IEGNDIQRLKGYPAIVFQHEIDHNLGVWFYDHDKDHPQLQ 175
Db 114 RDGKPFLEADGLLAICIQHEMDHLVGLKLFMDYLS---PLK 151

RESULT 15
US-09-188-820-4
; Sequence 4, Application US/09188820
; Patent No. 6190902
; GENERAL INFORMATION:
; APPLICANT: Bogosian, Gregg
; TITLE OF INVENTION: Deforylation of f-Met Peptides in
; Bacterial Expression Systems
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.
; ZIP: 77210
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/09/188,820
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/895,939
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Patterson, Melinda L.
; REGISTRATION NUMBER: 33,062
; REFERENCE/DOCKET NUMBER: MOPV:006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 787-1400
; TELEFAX: (713) 789-2679
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 169 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-09-188-820-4

Query Match 18.2%; Score 180.5; DB 3; Length 169;
Best Local Similarity 32.9%; Pred. No. 2.3e-12;
Matches 53; Conservative 31; Mismatches 54; Indels 23; Gaps 8;

Qy 22 LELPLTKKEKETLIAMREFLVNSQ-----DEEIAKRYGLRSGVGLAAPQINISKMTIAVL 76
Db 7 LHIP---DERLRKVKAPVEEVNAEIQRIVDDMFETMYA-BEGIGLAATQVDIHQRIIVI- 61
Qy 77 IPDDGSGKSYDYMVLNPKIVSHSVQEAYLPTG--EGCLSVDDNVAGLVHRHNKITIKAKD 134
Db 62 --DVSENRDERLVLINPELLEKSGE-----TGIEEGCLSIPEQRA-LVPRAEKVKIRALD 113
Qy 135 IEGNDIQRLKGYPAIVFQHEIDHNLGVWFYDHDKDHPQLQ 175
Db 114 RDGKPFLEADGLLAICIQHEMDHLVGLKLFMDYLS---PLK 151

```

Search completed: March 31, 2004, 16:37:32
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 16:36:46 ; Search time 41 Seconds
(without alignments)
1206.197 Million cell updates/sec

Title: US-09-896-580B-12

Perfect score: 991

Sequence: 1 MLTWKDIIRDGHPTLRQKAA.....KDHPLQHTDAVEVHQHHH 189

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1065169 seqs, 261661801 residues

Total number of hits satisfying chosen parameters: 1065169

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	991	100.0	213	10	US-09-896-580A-1
2	773	78.0	172	10	US-09-896-580A-7
3	446.5	45.1	203	9	US-09-862-005-2
4	386.5	39.0	155	9	US-09-862-005-4
5	228.5	23.1	216	10	US-09-896-580A-5
6	211.5	20.1	169	10	US-09-896-580A-3
7	199	20.1	267	15	US-10-359-513-4
8	199	20.1	284	15	US-10-359-513-10
9	188	19.0	277	15	US-10-353-513-9
10	184.5	18.6	160	10	US-09-896-580A-4
11	180.5	18.2	168	10	US-09-896-580A-2
12	180.5	18.2	169	14	US-10-189-505-1
13	179	18.1	170	10	US-09-896-580A-6
14	179	18.1	256	15	US-10-353-513-2
15	176	17.8	268	12	US-10-424-599-229783

16	169.5	17.1	204	15	US-10-289-762-4	Sequence 4, Appli
17	168	17.0	252	12	US-10-424-599-152628	Sequence 152628,
18	168	17.0	252	15	US-10-359-513-6	Sequence 6, Appli
19	168	17.0	264	12	US-10-425-114-52771	Sequence 52771, A
20	160.5	16.2	193	9	US-09-825-345-2	Sequence 2, Appli
21	160.5	16.2	193	9	US-09-738-626-6508	Sequence 6508, Ap
22	159	16.0	216	14	US-10-156-761-10570	Sequence 10570, A
23	152.5	15.4	138	10	US-09-896-580A-8	Sequence 8, Appli
24	151.5	15.3	174	10	US-09-882-227-118	Sequence 118, App
25	147.5	14.9	174	12	US-10-335-977-5438	Sequence 5438, Ap
26	147.5	14.9	177	12	US-10-335-977-5438	Sequence 5439, Ap
27	146.5	14.8	224	14	US-10-156-761-14878	Sequence 14878, A
28	143.5	14.5	221	15	US-10-359-513-8	Sequence 8, Appli
29	143	14.4	169	9	US-09-738-626-5271	Sequence 5271, Ap
30	143	14.4	186	12	US-10-425-114-61994	Sequence 61994, A
31	140	14.1	243	9	US-09-866-099-2	Sequence 2, Appli
32	140	14.1	243	14	US-10-149-256-6	Sequence 6, Appli
33	139	14.0	185	14	US-10-156-761-14655	Sequence 14655, A
34	98.5	9.9	210	14	US-10-149-256-4	Sequence 4, Appli
35	95	9.6	174	12	US-10-424-599-243598	Sequence 243598,
36	91	9.2	162	10	US-09-921-144-16	Sequence 16, Appl
37	91	9.2	163	10	US-09-921-144-14	Sequence 14, Appl
38	91	9.2	184	10	US-09-921-144-20	Sequence 20, Appl
39	91	9.2	185	10	US-09-921-144-18	Sequence 18, Appl
40	87.5	8.8	604	10	US-09-949-293-30	Sequence 30, Appl
41	87.5	8.8	604	15	US-10-260-937-26	Sequence 26, Appl
42	86.5	8.7	277	12	US-10-282-122A-45136	Sequence 45136, A
43	85	8.6	522	9	US-09-995-749A-11	Sequence 11, Appl
44	84.5	8.5	604	10	US-09-949-293-26	Sequence 26, Appl
45	84.5	8.5	604	15	US-10-260-937-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1
US-09-896-580A-1
; Sequence 1, Application US/09896580A
; Publication No. US20030170868A1
; GENERAL INFORMATION:
; APPLICANT: Baldwin, Eric
; TITLE OF INVENTION: CRYSTALLIZATION AND STRUCTURE OF STAPHYLOCOCCUS AUREUS PEPTIDE
; FILE REFERENCE: 268.6317.0101
; CURRENT APPLICATION NUMBER: US/09/896,580A
; PRIOR FILING DATE: 2002-08-19
; PRIOR FILING DATE: 2000-06-30
; PRIOR FILING DATE: 2000-06-30
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 1
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-896-580A-1

Query Match	100.0%	Score 991;	DB 10;	Length 213;
Best Local Similarity	100.0%	Pred. No. 1.5e-94;		
Matches 189;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MLTWKDIIRDGHPTLRQKAAELPLTKEKETLIAMREFLVNSQDEFIKRYGLRSVG 60		
Db	25	MLTWKDIIRDGHPTLRQKAAELPLTKEKETLIAMREFLVNSQDEFIKRYGLRSVG 84		
QY	61	LAAPQINISKMIIVLPDDGSGSYDVMVNPKVSHSVQEAYLPTGEGCLSDVDNAG 120		
Db	85	LAAPQINISKMIIVLPDDGSGSYDVMVNPKVSHSVQEAYLPTGEGCLSDVDNAG 144		
QY	121	LVRHKNITIKADIEGNDIQLRLKGYPAIVFQHEIDHLNGVMFYDHDKQHPLOPHDTA 180		


```

Db 145 LVHRHNKITIKAKDIEGNDIQLRLKGYPAIVFQHEIDLHNGVMFYDHDHDKDHPLOPHDTA 204
QY 181 VEVHQHHH 189
Db 205 VEVHQHHH 213

RESULT 2
US-09-896-580A-7
; Sequence 7, Application US/09896580A
; Publication No. US20030170868A1
; GENERAL INFORMATION:
; APPLICANT: Baldwin, Eric
; APPLICANT: Harris, Melissa
; TITLE OF INVENTION: CRYSTALLIZATION AND STRUCTURE OF STAPHYLOCOCCUS AUREUS PEPTIDE
; FILE REFERENCE: 268.6317 0101
; CURRENT APPLICATION NUMBER: US/09/896,580A
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: 60/215,555
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/215,550
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 7
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-896-580A-7

Query Match 78.0%; Score 773; DB 10; Length 172;
Best Local Similarity 90.3%; Pred. No. 4.7e-72;
Matches 168; Conservative 3; Mismatches 1; Indels 14; Gaps 10;

QY 1 MLTKMDIIRDGHPTRLQKAAELEPLTKKEKETLIAMREFLVNSQDEEIAKRYGLRSGVG 60
Db 1 MLTKMDIIRDGHPTRLQKAAELEPLTKKEKETLIAMREFLVNSQDEEIAKRYGLRSGVG 60
QY 61 LAAPQINISKRMIAVLIPDDGSGKSYD--LVNPKIVSHSVQAEYALPTGEGCLSDVDNVAG 120
Db 54 LAAPQINISKRMIAVLIPDDGSGKSYD--LVNPKIVSHSVQAEYALPTGEGCLSDVDNVAG 107
QY 121 LVHRHNKITIKAKDIEGNDIQLRLKGYPAIVFQHEIDLHNGVMFYDHDHDKDHPLOPHDTA 180
Db 108 LVHRHNRI-IKAKDIEGNDIQLRLKGYPAIVFQHEIDLHNGVMFYDHDHDKDHPLOPHDTA 166
QY 181 VEVHQH 186
Db 167 VEVHHH 172

RESULT 3
US-09-862-005-2
; Sequence 2, Application US/09862005
; Patent No. US20020058796A1
; GENERAL INFORMATION:
; APPLICANT: Lonetto, Michael A.
; TITLE OF INVENTION: def1
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: 21-May-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/991,023
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GW50010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2488

```

```

; APPLICATION NUMBER: US/09/862,005
; FILING DATE: 21-May-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/991,023
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GW50010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2488
; TELEFAX: 215-994-2222
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 203 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-862-005-2

Query Match 45.1%; Score 446.5; DB 9; Length 203;
Best Local Similarity 51.1%; Pred. No. 4.7e-38;
Matches 92; Conservative 35; Mismatches 46; Indels 7; Gaps 2;

QY 1 MLTKMDIIRDGHPTRLQKAAELEPLTKKEKETLIAMREFLVNSQDEEIAKRYGLRSGVG 60
Db 13 LIDMNDIIRREGNPLRTVAEEVTFPLSDQEIILGKMMQFLKHSQDPVMAEKMLRGVVG 72
QY 61 LAAPQINISKRMIAVLIPD-----DGSKGSYD--MLVNPKIIVSHSVQAEYALPTGEGCLS 113
Db 73 LAAPQLDISKRIITAVLVNIVEGETPQEAAYDLEAINMNPKIIVSHSVQDAALGEGCLS 132
QY 114 VDDNVAGLVHRHNKITIKAKDIEGNDIQLRLKGYPAIVFQHEIDLHNGVMFYDHDHDKHP 173
Db 133 VDRNPGYVVRHARVTVDYFDKGEKHKRIKLGKNSIVVQHEIDHNGIMFYDRINEKDP 192

RESULT 4
US-09-862-005-4
; Sequence 4, Application US/09862005
; Patent No. US20020058796A1
; GENERAL INFORMATION:
; APPLICANT: Lonetto, Michael A.
; TITLE OF INVENTION: def1
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/862,005
; FILING DATE: 21-May-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/991,023
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GW50010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2488

```

```

:
: TELEFAX: 215-994-2222
:
: TELEX: <UNKNOWN>
:
: INFORMATION FOR SEQ ID NO: 4:
:
:   SEQUENCE CHARACTERISTICS:
:
:     LENGTH: 155 amino acids
:
:     TYPE: amino acid
:
:     STRANDEDNESS: single
:
:     TOPOLOGY: linear
:
:   SEQUENCE DESCRIPTION: SEQ ID NO: 4:
: US-09-862-005-4

```

	Query Match	39.0%;	Score 386.5;	DB 9;	Length 155;
	Best Local Similarity	54.2%;	Pred. No. 5.4e-32;		
	Matches	78;	Conservative	28;	Mismatches 31; Indels 7; Gaps 2;
Qy	37	MRFLVNSDDEETAKRYGLRSGVGLAAPQINISKRMIAVLIPD-----DGSKGKSYDY--M	89		
Db	1	MMQFLKHSODPVMAEKMGRLGGVGLAAPOLDISKRIIAVLVPNIVERGETPQAYDLEAI	60		
Qy	90	LVNPKIVSHSVQAYLPTGGCGLSVDNDVAGLVHRHNKIIKAKDIGNDIQLRLKGYP	149		
Db	61	MYNPKIVSHSVQDAALGEGGCLSVDRNVGYVVRHARVIVDYEDKQEKHRIKLKGYNS	120		
Qy	150	IVFQHEIDHLGNVMFYDHIDKQHP	173		
Db	121	IVYQHEIDHNGIMFYDRINEKQD	144		

```

RESULT 5
US-09-896-580A-5
// Sequence 5, Application US/09896580A
// Publication NO. US2003017086A1
// GENERAL INFORMATION:
// APPLICANT: Baldwin, Eric
// APPLICANT: Harris, Melissa
// TITLE OF INVENTION: CRYSTALLIZATION AND ST
// TITLE OF INVENTION: DEFORMYLASE
// FILE REFERENCE: 268.6317 0101
// CURRENT APPLICATION NUMBER: US/09/896,580A
// CURRENT FILING DATE: 2002-08-19
// PRIOR APPLICATION NUMBER: 60/215,555
// PRIOR FILING DATE: 2000-06-30
// PRIOR APPLICATION NUMBER: 60/215,550
// PRIOR FILING DATE: 2000-06-30
// NUMBER OF SEQ ID NOS: 11
// SOFTWARE: PatentIn version 3.0
// SEQ ID NO 5
// LENGTH: 216
// TYPE: PR1
// ORGANISM: Mycoplasma pneumoniae
US-09-896-580A-5

```

Query Match	23.1%;	Score 228.5;	DB 10;	Length 216;
Best Local Similarity	32.2%;	Pred. No. 2.2e-15;		
Matches	56;	Conservative 36;	Mismatches 73;	Indels 9; Gaps 4;

QY	2	LTMKDIIRDCHPTLRQKAELPLPTKEKETLIAMREFLVNSQDEIIAKRYGLRSGVGL	61
DbB	33	LVLDVVKKEINPT-----KPVQFPDLQASLDLCIAKMAVVDASYNQD-AEKYGIIPGIGI	86
QY	62	AAPQINISKRMTAVLIPDGGSGSYDYMLVNPVKIVSHSVQEAFLPTGEGCLSYDDNVAGL	121
DbB	87	AAANQIGYWKQMFVHLMDD--GGVEHKLLINPKFINLSANKSFLKSGECLSPVQKHQY	144
QY	122	VIRHNKTIITKAND-TEGNDIQLRLKGYPAIVFQHEIDHLNGVMFYDHIDKHPL	174
DbB	145	VIRHEWITITGFDWLQOKEITTTATLFGMLQHFDFHLOGFYFYHRIINPL	198

RESULT 6
US-09-896-580A-3
; Sequence 3, Application US/09896580A
; Publication No. US20030170868A1

```

; GENERAL INFORMATION:
; APPLICANT: Baldwin, Eric
; APPLICANT: Harris, Melissa
; TITLE OF INVENTION: CRYSTALLIZATION AND ST
; TITLE OF INVENTION: CRYSTALLIZATION AND ST
; TITLE OF INVENTION: CRYSTALLIZATION AND ST
; FILE REFERENCE: 268 6317 0101
; CURRENT APPLICATION NUMBER: US/09/896,580A
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: 60/215,555
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/215,550
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
; US-09-896-580A-3

```

[illegible]

```

RESULT 7
US-10-359-513-4
  ; Sequence 4, Application US/10359513
  ; Publication No. US20030200559A1
  ; GENERAL INFORMATION:
  ; APPLICANT: Butler, Karlene
  ; APPLICANT: Falco, Karl
  ; APPLICANT: Guttridge, Steve
  ; APPLICANT: Harvell, Lealie T.
  ; TITLE OF INVENTION: PEPTIDE DEFORMYLASE
  ; FILE REFERENCE: BB1503 US NA
  ; CURRENT APPLICATION NUMBER: US/10/359,513
  ; CURRENT FILING DATE: 2003-02-06
  ; NUMBER OF SEQ ID NOS: 10
  ; SOFTWARE: Microsoft Office 97
  ; SEQ ID NO 4
  ; LENGTH: 267
  ; TYPE: PRT
  ; ORGANISM: Oryza sativa
US-10-359-513-4

```

[illegible]

RESULT 8
US-10-359-513-10
; Sequence 10, Application US/10359513
; Publication No. US20030200559A1
; GENERAL INFORMATION:
; APPLICANT: Butler, Karlene
; APPLICANT: Falco, Carl
; APPLICANT: Gutteridge, Steve
; APPLICANT: Harvell, Leslie T.
; TITLE OF INVENTION: PEPTIDE DEFORMYLASE
; FILE REFERENCE: BB1503 US NA
; CURRENT APPLICATION NUMBER: US/10/359,513
; CURRENT FILING DATE: 2003-02-06
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-359-513-10

Query Match 20.1%; Score 199; DB 15; Length 284;
Best Local Similarity 32.5%; Pred. No. 3.7e-12;
Matches 55; Conservative 34; Mismatches 54; Indels 26; Gaps 7;
QY 19 AAELRL-PLTKEEK--ETLIAMREFLVNSQDEEIAKRYGLRS-----GYCLA 62
Db 69 AADIQFEPPLKVKYKPPDILARKNRINTFDD-----NLRSLTDEMPVMYKTDGIGLS 122
QY 63 APQINISKRMIAVLIPDDGSGKSYDMLVNPVKIVSHSVQEAAYLPTGEGCLSVDDONVAGLV 122
Db 123 APQGVNVOLM-VENPAGVKGEGEIVLNP--VVKMSKLLVYEEGCLSPFGIYANVV 179
QY 123 HRHNKTIKADIEGNDIQLRLKGYPAIVFQHEIDHLNGVMFYDHIDK 171
Db 180 -RPDVKIDAQDVTGAKIKVKLSGLSARVFQHEFDHLQGLIFFORMSLD 227

RESULT 9
US-10-359-513-9
; Sequence 9, Application US/10359513
; Publication No. US20030200559A1
; GENERAL INFORMATION:
; APPLICANT: Butler, Karlene
; APPLICANT: Falco, Carl
; APPLICANT: Gutteridge, Steve
; APPLICANT: Harvell, Leslie T.
; TITLE OF INVENTION: PEPTIDE DEFORMYLASE
; FILE REFERENCE: BB1503 US NA
; CURRENT APPLICATION NUMBER: US/10/359,513
; CURRENT FILING DATE: 2003-02-06
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-10-359-513-9

Query Match 19.0%; Score 188; DB 15; Length 277;
Best Local Similarity 29.9%; Pred. No. 5e-11;
Matches 53; Conservative 30; Mismatches 66; Indels 28; Gaps 6;
QY 4 MKDIIRDGHPTLRQKAARELEPLTKEKETLIAMREFLVNSQDEEIAKRYGLRSGVGLAA 63
Db 88 MPDIIVKAGDVLHPSQDIPLEETGSR-----IQKIIEMVWVWENAPGVGLAA 137
QY 64 PQINISKRMIAV-----LIPDDGSGKSYD-----YMLVNPVKIVSHSVQEAAYLPTGE 109
Db 138 PQIGIPLKIIVLEDTNEIYSAPKDET-KAQRPRFGLLVINPLKKGKNTALF--FE 194

QY 110 GCLSVDDNVLGVHRHNKTIKADIEGNDIQLRLKGYPAIVFQHEIDHLNGVMFYD 166
Db 195 GCLSV-DGFRAVVERHLEVEVTGLDRNGKAIKVDASGQARILQHEYDHLDTGLYVD 250
RESULT 10
US-09-896-580A-4
; Sequence 4, Application US/09896580A
; Publication No. US20030170868A1
; GENERAL INFORMATION:
; APPLICANT: Baldwin, Eric
; APPLICANT: Harris, Melissa
; TITLE OF INVENTION: CRYSTALLIZATION AND STRUCTURE OF STAPHYLOCOCCUS AUREUS PEPTIDE
; TITLE OF INVENTION: DEFORMYLASE
; FILE REFERENCE: 268.6317 0101
; CURRENT APPLICATION NUMBER: US/09/896,580A
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: 60/215,555
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/215,550
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-09-896-580A-4
Query Match 18.6%; Score 184.5; DB 10; Length 160;
Best Local Similarity 35.1%; Pred. No. 5.3e-11;
Matches 60; Conservative 27; Mismatches 59; Indels 25; Gaps 8;
QY 2 LTKMDIIRDGHPT-LRQKAARELEPLTKEKETLIAMREFLVNSQDEEIAKRYGLRSGVG 60
Db 1 MAVKKVVT--HPAVLETPAETVTVFDKKLKLDDMDYDTMLE-----MDGVG 46
QY 61 LAAPQINISKRMIAVLIPDDGSGKSYDMLVNPVKIVSHSVQEAAYLPTG-EGCLSVDDNVA 119
Db 47 LAAPQIGILKRAAVVEIGDD-RGR---IDLVPFELKSGEQ-----TGIEGCLSF-PNVY 97
QY 120 GLVHRHNKTIKADIEGNDIQLRLKGYPAIVFQHEIDHLNGVMFYDHIDK 170
Db 98 GDVTRADYVKVRAFNRQCKPFILEARGFLARAVQHMDHLDGLVLTFSKISK 148
RESULT 11
US-09-896-580A-2
; Sequence 2, Application US/09896580A
; Publication No. US20030170868A1
; GENERAL INFORMATION:
; APPLICANT: Baldwin, Eric
; APPLICANT: Harris, Melissa
; TITLE OF INVENTION: CRYSTALLIZATION AND STRUCTURE OF STAPHYLOCOCCUS AUREUS PEPTIDE
; TITLE OF INVENTION: DEFORMYLASE
; FILE REFERENCE: 268.6317 0101
; CURRENT APPLICATION NUMBER: US/09/896,580A
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: 60/215,555
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/215,550
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-896-580A-2
Query Match 18.2%; Score 180.5; DB 10; Length 168;
Best Local Similarity 32.9%; Pred. No. 1.5e-10;
Matches 53; Conservative 31; Mismatches 54; Indels 23; Gaps 8;


```

; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 229783
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(268)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MET3847_49518C.1.pap
US-10-424-599-229783

```

```

Query Match      17.8%; Score 176; DB 12; Length 268;
Best Local Similarity 29.2%; Pred. No. 8.4e-10;
Matches 45; Conservative 31; Mismatches 74; Indels 4; Gaps 4;

Qy 18 KAAELEPLTKEEKETLIAMREFLVNSQDESIKRYGLRSGVGLAAPQINISKRMIAVLI 77
Db 76 KIVEYDPLRLARNKRIVAFDDSLKLVHEMFDMY-KTDGIGLSAPQLGINVQLM-VFN 133
Qy 78 PDDGSGKSYDMLNPKIVSHSVQEAYLPTGEGCLSVDDNVAGLVHRHNKITIKAKDIEG 137
Db 134 PVGERGEGERIVLVNPRVSYSKKLTLENEG-XCLSF-PGINADVCRPESVKIDARDING 191
Qy 138 NDIQLRLKGYPAIVFQHEIDHNGVMFYDHDKD 171
Db 192 TRFSVNLSDLPARIFQHEFDHLQGLFFERMTEE 225

```

Search completed: March 31, 2004, 16:42:11
Job time : 46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 16:10:29 ; Search time 58 Seconds
(without alignments)
920.715 Million cell updates/sec

Title: US-09-896-580B-12

Perfect score: 991

Sequence: 1 MLTWKDIIRDGHPTLRQAA.....KDHPLQHTDAVEVHQHHH 189

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	991	100.0	213	5	AAM48343
2	943	95.2	207	6	ABM73276 Staphyloc
3	938	94.7	183	2	AAM483186 Staphyloc
4	795	80.2	183	4	AAG81798 S. epider
5	795	80.2	210	5	ABP38601 Staphyloc
6	754	76.1	150	2	ABP38601 Staphyloc
7	545	55.0	146	4	AAG82977 S. epider
8	533.5	53.8	183	5	ABBA7633 Staphyloc
9	498.5	50.3	187	6	ABD10150 Alloioococ
10	498.5	50.3	193	6	ADB10148 Alloioococ
11	466	47.0	204	5	ABP26597 Streptoco
12	453	45.7	204	5	ABP26597 Streptoco
13	450.5	45.5	203	6	ABU01906 Streptoco
14	446.5	45.1	203	2	AAW77215 Streptoco
15	426	43.0	211	5	ABBS3869 Streptoco
16	386.5	39.0	155	2	AAW77216 Streptoco
17	386.5	39.0	155	2	AAW77216 Streptoco
18	228.5	23.1	169	5	AAM48347 Mycoplasma
19	211.5	21.3	169	5	AAM48347 Mycoplasma
20	209.5	21.1	176	7	ADC96995 E. faeciu
21	203.5	20.5	273	3	AAG20819 Arabidops
22	199	20.1	267	7	ABR63183 Rice pept
23	199	20.1	284	6	ABR63183 Rice pept
24	198.5	20.0	164	6	ADA34567 Acinetoba
25	198	20.0	181	2	AAY04475 Chlamydia

26	197	19.9	158	3	AAG20821 Arabidops
27	197	19.9	162	3	AAG20820 Arabidops
28	188	19.0	277	7	ABR63184 Tomato pe
29	186.5	18.8	160	2	AAY16109 A formate
30	184.5	18.6	160	5	AAM48346 Bacillus
31	183	18.5	259	3	AAG28425 Arabidops
32	183	18.5	259	3	AAG28425 Arabidops
33	183	18.5	264	3	AAG28424 Arabidops
34	183	18.5	269	3	AAG28459 Arabidops
35	180.5	18.2	168	5	AAM48344 Escherich
36	180.5	18.2	169	2	AAW47032 E. coli pe
37	180.5	18.2	169	6	ABP57095 Escherich
38	179	18.1	170	5	AAM48348 Staphyloc
39	179	18.1	256	7	ABR63180 Maize pep
40	176	17.8	168	7	ADD35426 Pseudomon
41	176	17.8	203	6	ADA33539 Acinetoba
42	175	17.7	168	7	ADD35424 Pseudomon
43	175	17.7	171	6	ABM69391 Phototrab
44	174.5	17.6	166	2	RAY36827 Protein 1
45	169.5	17.1	204	2	AAY34586 C. pneumo

ALIGNMENTS

RESULT 1

AAM48343

ID AAM48343 standard; protein; 213 AA.

AC

XX AAM48343;

XX

DT 23-APR-2002 (first entry)

XX

DE Staphylococcus aureus peptide deformylase with C-terminal 6xHis tag.

XX

KW Peptide deformylase; enzyme; protein coordinate data; antibacterial; pdf.

XX

OS Staphylococcus aureus.

XX

PN WO200202758-A2.

XX

PD 10-JAN-2002.

XX

PF 29-JUN-2001; 2001WO-US020777.

XX

PR 30-JUN-2000; 2000US-0215550P.

XX

PA (PHAA) PHARMACIA & UPJOHN CO.

XX

PI Baldwin ET, Harris MS;

XX

DR WPI; 2002-148012/19.

XX

PT Crystalline Staphylococcus aureus peptide deformylase useful for solving structures of other molecules or molecular complexes, and designing modifiers of peptide deformylase activity.

PT

XX Claim 67; Fig 3; 149pp; English.

PS

CC The present sequence is Staphylococcus aureus peptide deformylase (pdf). The present invention relates to the crystal structure for pdf, which can be used in combination with a computer-assisted method for identifying, designing and making a potential modifier of S.aureus pdf activity.

XX

CC Modifiers of pdf are useful for blocking bacterial growth

XX

SQ Sequence 213 AA;

Query Match

Best Local Similarity 100.0%; Score 991; DB 5; Length 213;

Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 MLTWKDIIRDGHPTLRQAAELEPLTKEETLIAMEFLVNSQDEIAKRYGLRSGVG 60
|||||

Db 25 MLTKMDIIRDGHTLRQKAAELEPLTKKEKETLIAMREFLVNSQDEEIAKRYGLRSGVG 84
 QY 61 LAAPQINISKRMIANLIPDDGSGKSYDYMVNPKIVSHSVQEAYLPTGEGCLSVDDNVAG 120
 Db 85 LAAPQINISKRMIANLIPDDGSGKSYDYMVNPKIVSHSVQEAYLPTGEGCLSVDDNVAG 144
 QY 121 LVHRNKITIKAKDIEGNDIQRLKGYPAIVFQHEIDHLNGVMFYDHDKDHPLQPHDTA 180
 Db 145 LVHRNKITIKAKDIEGNDIQRLKGYPAIVFQHEIDHLNGVMFYDHDKDHPLQPHDTA 204
 QY 181 VEVQHSHH 189
 Db 205 VEVQHSHH 213

RESULT 2
 ABW73276
 ID ABW73276 standard; protein; 207 AA.

XX AC ABW73276;
 XX DT 20-NOV-2003 (first entry)
 XX DE Staphylococcus aureus protein #2516.
 XX KW Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
 XX KW enzymatic assay; antibiotic target.
 XX OS Staphylococcus aureus.
 XX PN WO200294868-A2.
 XX PD 28-NOV-2002.
 XX PF 27-MAR-2002; 2002WO-IB002637.
 XX PR 27-MAR-2001; 2001GB-00007661.
 XX PA (CHIR-) CHIRON SPA.
 XX PI Masignani V, Mora M, Scarselli M;
 XX WPI; 2003-120786/11.
 XX DR N-PSDB; ACF74836.
 XX PT New Staphylococcus aureus protein, useful as a vaccine for treating or
 XX PT preventing Staphylococcal infection, specifically an infection caused by
 XX PT S. aureus, e.g. sepsis.

XX PS Claim 1; SEQ ID NO 5032; 49pp; English.

XX CC The invention relates to novel genes and encoded proteins from
 XX CC Staphylococcus aureus. A composition comprising the S. aureus protein, a
 XX CC nucleic acid encoding the protein, or an antibody to the protein, is
 XX CC useful as a pharmaceutical, particularly as a vaccine for treating or
 XX CC preventing infection due to Staphylococcus bacteria, specifically an
 XX CC infection caused by S. aureus. The composition is particularly useful for
 XX CC treating or preventing sepsis in a patient. The composition can also be
 XX CC used for diagnostics. The protein is also used in an assay for enzymatic
 XX CC studies and as a target for antibiotics. This sequence represents one of
 XX CC the novel S. aureus proteins of the invention

XX SQ Sequence 207 AA;

Query Match 95.2%; Score 943; DB 6; Length 207;
 Best Local Similarity 99.5%; Pred. No. 1.5e-95;
 Matches 182; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLTKMDIIRDGHTLRQKAAELEPLTKKEKETLIAMREFLVNSQDEEIAKRYGLRSGVG 60
 Db 25 MLTKMDIIRDGHTLRQKAAELEPLTKKEKETLIAMREFLVNSQDEEIAKRYGLRSGVG 84
 QY 61 LAAPQINISKRMIANLIPDDGSGKSYDYMVNPKIVSHSVQEAYLPTGEGCLSVDDNVAG 120

Db 85 LAAPQINISKRMIANLIPDDGSGKSYDYMVNPKIVSHSVQEAYLPTGEGCLSVDDNVAG 144
 QY 121 LVHRNKITIKAKDIEGNDIQRLKGYPAIVFQHEIDHLNGVMFYDHDKDHPLQPHDTA 180
 Db 145 LVHRNKITIKAKDIEGNDIQRLKGYPAIVFQHEIDHLNGVMFYDHDKDHPLQPHDTA 204
 QY 181 VEV 183
 Db 205 VEV 207

RESULT 3

AAW83186
 ID AAW83186 standard; protein; 193 AA.

XX AC AAW83186;
 XX DT 11-FEB-1999 (first entry)
 XX DE Staphylococcus aureus deformylase 1.
 XX KW Staphylococcus aureus; Def1; deformylase 1; bacterial infections;
 XX KW antimicrobial screening; diagnosis; antibiotic.
 XX OS Staphylococcus aureus.
 XX PN EP879879-A2.
 XX PD 25-NOV-1998.
 XX PF 21-MAY-1998; 98EP-00304076.
 XX PR 21-MAY-1997; 97US-0048706P.
 XX PR 15-AUG-1997; 97US-00911844.
 XX PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX PI Lonetto MA, Sylvester DR, Warren RL;
 XX DR WPI; 1998-596871/51.
 XX DR N-PSDB; AAV70274.
 XX PT New nucleic acid encoding polypeptide deformylase of Staphylococcus
 XX PT aureus - useful for treatment, prevention and diagnosis of bacterial
 XX PT infections, and for antimicrobial screening.

XX PS Claim 12; Page 19-20; 23pp; English.

XX CC The present sequence represents deformylase 1 (Def1) isolated from
 XX CC Staphylococcus aureus WCUH 29 (NCIMB 40771). Host cells containing
 XX CC vectors comprising nucleic acid molecules encoding Def1 are used to
 XX CC produce recombinant Def1 which have polypeptide deformylase (Def)
 XX CC activity. Def1 proteins are used to treat conditions requiring Def
 XX CC activity, while Def1 antagonists, e.g. Ab, are used to treat conditions
 XX CC requiring reduction in Def activity, especially they are antibacterials
 XX CC for treating a wide range of infections caused by Staphylococcus, also by
 XX CC some other bacteria, e.g. Helicobacter pylori. Def1 proteins are also
 XX CC used to screen for compounds that interact specifically with it (i.e.
 XX CC potential antibacterials); to study the role of Def in disease; for
 XX CC generating Ab and in protective vaccines (to generate an antibody and/or
 XX CC T-cell response). Vaccination may also be with a vector that contains a
 XX CC nucleic acid molecule encoding Def1. Ab are used therapeutically and for
 XX CC diagnosing in standard immunoassays. Antagonists may inhibit binding of
 XX CC bacteria to extracellular matrix proteins and to in-dwelling devices, or
 XX CC they inhibit normal progression of infection. Fragments of nucleic acid
 XX CC molecules encoding Def1 are useful as hybridisation probes and/or
 XX CC amplification primers for isolation of full-length Def-encoding sequences
 XX CC or related genes; to detect expression in infected tissue, e.g. for
 XX CC diagnosis and staging, and to identify mutations (e.g. for serotyping)

XX SQ Sequence 183 AA;

Query Match 94.7%; Score 938; DB 2; Length 183;
 Best Local Similarity 98.9%; Pred. No. 4.6e-95;
 Matches 181; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLTWKDIIRDGHPTLRQAAEELPLTKKEKETLAMEEFLVNSODEEIAKRYGLRSGVG 60
 DB 1 MLTWKDIIRDGHPTLRQAAEELPLTKKEKETLAMEEFLVNSODEEIAKRYGLRSGVG 60
 QY 61 LAAPQINISKMIAMVLIIPDDGSGKSYDMLVNPVKIVSHSVQEAYLPTGEGCLSVDDNVAG 120
 DB 61 LAAPQINISKMIAMVLIIPDDGSGKSYDMLVNPVKIVSHSVQEAYLPTGEGCLSVDDNVAG 120
 QY 121 LVHRHNKTIKAKDIEGNDIQRLKGYPAIVFOHEIDHLNGVMFYDHDKHPLQPHDTA 180
 DB 121 LVHRHNKTIKAKDIEGNDIQRLKGYPAIVFOHEIDHLNGVMFYDHDKHPLQPHDTA 180
 QY 181 VEV 183
 DB 181 VEV 183

RESULT 4

AAG81798
 ID AAG81798 standard; protein; 183 AA.

AC AAG81798;
 XX

DT 03-SEP-2001 (first entry)

XX S. epidermidis open reading frame protein sequence SEQ ID NO:690.

DE Staphylococcus epidermidis SR1 strain; infection; diagnosis; vaccination;
 KW endocarditis.

XX Staphylococcus epidermidis.

XX WO200134809-A2.

XX 17-MAY-2001.

XX 09-NOV-2000; 2000WO-US030782.

XX 09-NOV-1999; 99US-0164258P.

XX (GLAXO) GLAXO GROUP LTD.

XX Kimmerly WJ;

XX WPI; 2001-316495/33.

XX N-PSDB; AAH52648.

XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
 PT useful for vaccinating against infections, e.g. endocarditis.

XX Claim 18; Page 217; 2188pp; English.

XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
 CC (II), given in AAG81454 to AAG81120, from Staphylococcus epidermidis. (I)
 CC and (II) can have antibacterial activity and therefore can be used in
 CC vaccination. The nucleic acids (II) may be used to produce the S.
 CC epidermidis polypeptides (II) via the production of vectors containing
 CC them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
 CC polynucleotide sequences from the present invention. AAH55091 to AAH55098
 CC represent oligonucleotide sequences and primers which are used in the
 CC exemplification of the present invention. N.B. The present invention
 CC specifically claims all the polynucleotide sequences given in the
 CC sequence listing of the present specification, however the sequence

CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
 CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
 CC for SEQ ID NO:4455 to 4464

XX Sequence 183 AA;

Query Match 80.2%; Score 795; DB 4; Length 183;
 Best Local Similarity 79.8%; Pred. No. 2.9e-79;
 Matches 146; Conservative 23; Mismatches 14; Indels 0; Gaps 0;

QY 1 MLTWKDIIRDGHPTLRQAAEELPLTKKEKETLAMEEFLVNSODEEIAKRYGLRSGVG 60
 DB 1 MLTWKDIIRDGHPTLRQAAEELPLTKKEKETLAMEEFLVNSODEEIAKRYGLRSGVG 60
 QY 61 LAAPQINISKMIAMVLIIPDDGSGKSYDMLVNPVKIVSHSVQEAYLPTGEGCLSVDDNVAG 120
 DB 61 LAAPQINISKMIAMVLIIPDDGSGKSYDMLVNPVKIVSHSVQEAYLPTGEGCLSVDDNVAG 120
 QY 121 LVHRHNKTIKAKDIEGNDIQRLKGYPAIVFOHEIDHLNGVMFYDHDKHPLQPHDTA 180
 DB 121 LVHRHNKTIKAKDIEGNDIQRLKGYPAIVFOHEIDHLNGVMFYDHDKHPLQPHDTA 180
 QY 181 VEV 183
 DB 181 VEV 183

RESULT 5

ABP38601
 ID ABP38601 standard; protein; 210 AA.

XX ABP38601;

XX 24-JUL-2002 (first entry)

XX Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3446.

XX Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
 KW antibacterial; gene therapy.

XX Staphylococcus epidermidis.

XX US6380370-B1.

XX 30-APR-2002.

XX 13-AUG-1998; 98US-00134001.

XX 14-AUG-1997; 97US-0055779P.

XX 08-NOV-1997; 97US-0064964P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Doucette-Stamm LA, Bush D;

XX WPI; 2002-381255/41.

XX N-PSDB; ABN91146.

XX Novel isolated nucleic acid encoding a Staphylococcus epidermis
 PT polypeptide, useful for diagnosing and treating bacterial infections.

XX Disclosure; SEQ ID NO 3446; 267pp; English.

XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
 CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
 CC given in ABP3124 to ABP3960. The S. epidermidis sequences have
 CC antibacterial activity and can be used in gene therapy. The sequences can
 CC also be used in the diagnosis and treatment of bacterial infections,
 CC particularly S. epidermidis infections. The sequences can be used to
 CC screen for compounds able to interfere with the S. epidermidis life cycle
 CC or inhibit S. epidermidis infection. N.B. The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the USPTO web site

XX SQ Sequence 210 AA;

Query Match 80.2%; Score 795; DB 5; Length 210;
 Best Local Similarity 79.8%; Pred. No. 3.6e-79;
 Matches 146; Conservative 23; Mismatches 14; Indels 0; Gaps 0;

QY 1 MLTMDIIRDGHPTLRQKAAELPLTKKEETLIAMREFLVNSQDEETAKYGLRSGVG 60
 Db 28 MITMDIIRDGHPTLRQKAAELPLTKKEETLIAMREFLVNSQDEETAKYGLRSGVG 87

QY 61 LAAPQINISKRMIAYLIPDDGSGKSYDYMVNPKIVSHSVQAYLPTGEGCLSVDDNVAG 120
 Db 88 LAAPQINISKRMIAYLIPDDGSGKSYDYMVNPKIVSHSVQAYLPTGEGCLSVDDNVAG 147

QY 121 LVHRHNKITTAKDIEGNDIQRLKGYPAIVFQHEIDHLNGVMFVDHIDKHPLOPHYDA 180
 Db 148 LVHRHRTVITKAQDIDGNDVRLKGYPAIVFQHEIDHLNGIMFYDIDANEPLKPHBEA 207

QY 181 VEV 183
 Db 208 VEV 210

RESULT 6
 AAW83187
 ID AAW83187 standard; protein; 150 AA.
 AC AAW83187;
 DT 11-FEB-1999 (first entry)
 XX XX
 DE Staphylococcus aureus deformylase 1 ORF protein sequence.
 XX Staphylococcus aureus; Defl; deformylase 1; bacterial infections;
 KW antimicrobial screening; diagnosis; antibacterial.
 XX Staphylococcus aureus.
 XX EP879879-A2.
 XX 25-NOV-1998.
 XX 21-MAY-1998; 98EP-00304076.
 XX 21-MAY-1997; 97US-0048706P.
 XX 15-AUG-1997; 97US-00911844.
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 XX Lonetto MA, Sylvester DR, Warren RL;
 XX WPI; 1998-596871/51.
 XX N-PSDB; AAV70275.
 XX New nucleic acid encoding polypeptide deformylase of Staphylococcus aureus - useful for treatment, prevention and diagnosis of bacterial infections, and for antimicrobial screening.
 XX Disclosure; Page 21; 23pp; English.

CC The present sequence represents deformylase 1 (Defl) isolated from Staphylococcus aureus WCUH 29 (NCIMB 40771). Host cells containing vectors comprising nucleic acid molecules encoding Defl are used to produce recombinant Defl which have polypeptide deformylase (Def) activity. Defl proteins are used to treat conditions requiring Def activity, while Defl antagonists, e.g. Ab, are used to treat conditions requiring reduction in Def activity, especially they are antibacterials for treating a wide range of infections caused by Staphylococcus, also by some other bacteria, e.g. Helicobacter pylori. Defl proteins are also used to screen for compounds that interact specifically with it (i.e. potential antibacterials); to study the role of Def in disease; for generating Ab and in protective vaccines (to generate an antibody and/or

CC T-cell response). Vaccination may also be with a vector that contains a nucleic acid molecule encoding Defl. Ab are used therapeutically and for diagnosing in standard immunoassays. Antagonists may inhibit binding of bacteria to extracellular matrix proteins and to in-dwelling devices, or they inhibit normal progression of infection. Fragments of nucleic acid molecules encoding Defl are useful as hybridisation probes and/or amplification primers for isolation of full-length Def-encoding sequences or related genes; to detect expression in infected tissue, e.g. for diagnosis and staging, and to identify mutations (e.g. for serotyping)

XX SQ Sequence 150 AA;

Query Match 76.1%; Score 754; DB 2; Length 150;
 Best Local Similarity 98.7%; Pred. No. 7.4e-75;
 Matches 148; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLTMDIIRDGHPTLRQKAAELPLTKKEETLIAMREFLVNSQDEETAKYGLRSGVG 60
 Db 1 MLTMDIIRDGHPTLRQKAAELPLTKKEETLIAMREFLVNSQDEETAKYGLRSGVG 60

QY 61 LAAPQINISKRMIAYLIPDDGSGKSYDYMVNPKIVSHSVQAYLPTGEGCLSVDDNVAG 120
 Db 61 LAAPQINISKRMIAYLIPDDGSGKSYDYMVNPKIVSHSVQAYLPTGEGCLSVDDNVAG 120

QY 121 LVHRHNKITTAKDIEGNDIQRLKGYPAI 150
 Db 121 LVHRHNKITTAKDIEGNDIQRLKGYPAI 150

RESULT 7
 AAG82977
 ID AAG82977 standard; protein; 146 AA.
 AC AAG82977;
 DT 03-SEP-2001 (first entry)
 XX XX
 DE S. epidermidis open reading frame protein sequence SEQ ID NO:3048.
 XX Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination;
 KW endocarditis.
 XX Staphylococcus epidermidis.
 XX WO200134809-A2.
 XX 17-MAY-2001.
 XX 09-NOV-2000; 2000WO-US030782.
 XX 09-NOV-1999; 99US-0164258P.
 XX (GLAX) GLAXO GROUP LTD.
 XX Kimmerly WJ;
 XX WPI; 2001-316495/33.
 XX N-PSDB; AAH53827.
 XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis, useful for vaccinating against infections, e.g. endocarditis.
 XX Claim 18; Page 802; 2188pp; English.
 XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides (II), given in AAH51454 to AAG83120, from Staphylococcus epidermidis. (I) and (II) can have antibacterial activity and therefore can be used in vaccination. The nucleic acids (I) may be used to produce the S. epidermidis polypeptides (II) via the production of vectors containing them which are used to produce hosts cells which express the polypeptides. The polypeptides (II) (and/or nucleic acids) may then be used to vaccinate subjects and to raise antibodies against the bacteria. The polypeptides may also be used to assay for other inhibitors of their

CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
 CC polynucleotide sequences from the present invention. AAH55091 to AAH55098
 CC represent oligonucleotide sequences and primers which are used in the
 CC exemplification of the present invention. N.B. The present invention
 CC specifically claims all the polynucleotide sequences given in the
 CC sequence listing of the present specification, however the sequence
 CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
 CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
 CC for SEQ ID NO:4455 to 4464
 XX
 XX Sequence 146 AA;

Query Match 55.0%; Score 545; DB 4; Length 146;
 Best Local Similarity 81.7%; Pred. No. 8.8e-52;
 Matches 103; Conservative 12; Mismatches 11; Indels 0; Gaps 0;
 QY 1 MLTWKDIIRDGHPTLRQKAAELPLTKKEKETLIAMREFLVNSQDEEIAKRYGLRSGVG 60
 DB 1 MITWMDIIRDGHPTLRQKAAELPLTKKEKETLIAMREFLVNSQDEEIAKRYGLRSGVG 60
 QY 61 LAAPQINISKRMIAVLIPDDGSGKSYDMLVNPKIVSHSVQAEYLPTEGCLSVDDNVAG 120
 DB 61 LAAPQINISKRMIAVLIPDDGSGKSYDMLVNPKIVSHSVQAEYLPTEGCLSVDDNVAG 120
 QY 121 LVHRHN 126
 DB 121 LVHRHN 126

RESULT 8
 ABB47633
 ID ABB47633 standard; protein; 183 AA.
 XX
 AC ABB47633;
 DT 05-FEB-2002 (first entry)
 XX
 DE Listeria monocytogenes protein #337.
 KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 KW vitamin B12; bacterial infection; disease.
 XX
 OS Listeria monocytogenes.
 XX
 PN WO200177335-A2.
 XX
 PD 18-OCT-2001.
 XX
 PF 11-APR-2001; 2001WO-FR001118.
 XX
 PR 11-APR-2000; 2000FR-00004629.
 XX
 PA (INSP) INST PASTEUR.

XX Buchrieser C, Frangeul L, Couve E, Rueniok C, Psihi H, Dehoux P;
 PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
 PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
 PI Dominguez-Bernal G, Garrido-Garcia P, Tietz-Martinez A, Amend A;
 PI Chakraborty T, Dommann E, Hain T, Berche P, Charbit A, Durant L;
 PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
 PI Maduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
 PI Rose M, Voss H;
 XX
 XX WPI; 2002-010914/01.
 DR
 XX
 XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
 PT and prevention of Listeria and related bacterial infections, and related
 PT polypeptides.
 XX
 PS Claim 6; SEQ ID NO 338; 192pp; French.
 XX

CC The present invention relates to the genome sequence of Listeria
 CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in L.
 CC monocytogenes and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies, identification of L. monocytogenes and related organisms, and
 CC for biosynthesis and biodegradation, especially biosynthesis of vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate L. monocytogenes-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccine compositions for the treatment or prevention of infections by L.
 CC monocytogenes and related organisms. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 183 AA;

Query Match 53.8%; Score 533.5; DB 5; Length 183;
 Best Local Similarity 57.5%; Pred. No. 2.3e-50;
 Matches 104; Conservative 30; Mismatches 44; Indels 3; Gaps 2;
 QY 1 MLTWKDIIRDGHPTLRQKAAELPLTKKEKETLIAMREFLVNSQDEEIAKRYGLRSGVG 60
 DB 1 MLTWMDIIRDGHPTLRQKAAELPLTKKEKETLIAMREFLVNSQDEEIAKRYGLRSGVG 60
 QY 61 LAAPQINISKRMIAVLIPDDGSGKSYDMLVNPKIVSHSVQAEYLPTEGCLSVDDNVAG 120
 DB 61 LAAPQINISKRMIAVLIPDDGSGKSYDMLVNPKIVSHSVQAEYLPTEGCLSVDDNVAG 120
 QY 121 LVHRHNKTIKAKDIEGNDIQLKGYPAIVFOHEIDHLNGVMFYDHDKQHP--LQPH 178
 DB 120 YVVRSEVTDADFENGTPKLPFGYPAIVFOHEIDHLNGVMFYDHDKQHP--LQPH 178
 QY 179 D 179
 DB 180 D 180

RESULT 9
 ADB10150
 ID ADB10150 standard; protein; 187 AA.
 XX
 AC ADB10150;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Alloicoccus otitis antigenic protein SEQ ID NO:4910.
 XX
 KW Alloicoccus otitis; antigenic protein; immunogenic; immunisation;
 KW gene therapy; Gram-positive bacterium; infection.
 XX
 OS Alloicoccus otitis.
 XX
 PN WO2003048304-A2.
 XX
 PD 12-JUN-2003.
 XX
 PF 25-NOV-2002; 2002WO-US036123.
 XX
 PR 29-NOV-2001; 2001US-0333777P.
 PR 18-NOV-2002; 2002US-0426742P.
 XX
 PA (AMHP) WYETH HOLDINGS CORP.
 XX
 PI Fletcher LD, Mcmichael JC, Russell DP, Zagursky RJ;
 XX
 DR WPI; 2003-505284/47.
 DR N-PSDB; ADB10153.
 XX

KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 XX antiinflammatory; infection; vaccine; meningitis; gene therapy.
 OS Streptococcus pyogenes.

PN WO200234771-A2.

PD 02-MAY-2002.

XX 29-OCT-2001; 2001WO-GB004789.

XX 27-OCT-2000; 2000GB-00026333.

PR 24-NOV-2000; 2000GB-00028727.

PR 07-MAR-2001; 2001GB-00005640.

XX (CHIR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.

XX Telford J, Massignani V, Margarit Y Rosl, Grandi G, Fraser C;

PI Tettelin H;

XX WPI; 2002-352536/38.

DR N-PSDB; ABN67228.

XX New Streptococcus protein for the treatment or prevention of infection or

PT disease caused by Streptococcus bacteria, such as meningitis, and for

PT detecting a compound that binds to the protein.

XX Claim 1; Page 3389; 4525pp; English.

XX The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.

CC Nucleic acids encoding (I) are used to detect Streptococcus in a

CC biological sample. (I) is used to determine whether a compound binds to

CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be

CC used as a vaccine or diagnostic composition. The disease caused by

CC Streptococcus that is prevented or treated may be meningitis. Nucleic

CC acid encoding (I) may be used to recombinantly produce (I) and may be

CC used in gene therapy. Antibodies to (I) are used for affinity

CC chromatography, immunoassays, and distinguishing/identifying

CC Streptococcus proteins

XX Sequence 204 AA;

SQ Query Match 47.0%; Score 466; DB 5; Length 204;

Best Local Similarity 53.6%; Pred. No. 7.8e-43;

Matches 98; Conservative 31; Mismatches 46; Indels 8; Gaps 2;

QY 1 MLTMDIIRGDHPTLRQAAELELPITKEEKTLLAMREFLVNSQDEETAKRYGLRSGVG 60

DB 13 LITMDIIRGNTLRQAAELELPITKEEKTLLAMREFLVNSQDEETAKRYGLRSGVG 72

QY 61 LAAPQINISKRMIAVLIPD--DGSCK-----SYDMLVNPNTVSHSVQEAYLPTGEGCL 112

DB 73 LAAPQIDVSKRIITAVLPNLPDKEGPNPKPEAYSWQEVLPNPKVSHSVQDAALSDGEGCL 132

QY 113 SVDDNVAGLVHRHNNKTIKAKDIEGNDIQLRLKGYPAIVFQHEIDHLNGVMFYDHDKDH 172

DB 133 SVDRVVEGYVVRHRTVDYDKGQHQRIKLKGYNAIVVQHEIDHNGVLFYDRINAKN 192

QY 173 PLQ 175

DB 193 PFE 195

RESULT 12

ABP26596

ID ABP26596 standard; protein; 204 AA.

XX AC ABP26596;

XX DT 02-JUL-2002 (first entry)

XX Streptococcus polypeptide SEQ ID NO 2368.

XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;

XX group A streptococcus; Streptococcus pyogenes; antibacterial;

XX antiinflammatory; infection; vaccine; meningitis; gene therapy.

OS Streptococcus agalactiae.

XX WO200234771-A2.

XX 02-MAY-2002.

XX 29-OCT-2001; 2001WO-GB004789.

XX 27-OCT-2000; 2000GB-00026333.

PR 24-NOV-2000; 2000GB-00028727.

PR 07-MAR-2001; 2001GB-00005640.

XX (CHIR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.

XX Telford J, Massignani V, Margarit Y Rosl, Grandi G, Fraser C;

PI Tettelin H;

XX WPI; 2002-352536/38.

DR N-PSDB; ABN67227.

XX New Streptococcus protein for the treatment or prevention of infection or

PT disease caused by Streptococcus bacteria, such as meningitis, and for

PT detecting a compound that binds to the protein.

XX Claim 1; Page 3389; 4525pp; English.

XX The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by

CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.

CC Nucleic acids encoding (I) are used to detect Streptococcus in a

CC biological sample. (I) is used to determine whether a compound binds to

CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be

CC used as a vaccine or diagnostic composition. The disease caused by

CC Streptococcus that is prevented or treated may be meningitis. Nucleic

CC acid encoding (I) may be used to recombinantly produce (I) and may be

CC used in gene therapy. Antibodies to (I) are used for affinity

CC chromatography, immunoassays, and distinguishing/identifying

CC Streptococcus proteins

XX Sequence 204 AA;

SQ Query Match 45.7%; Score 453; DB 5; Length 204;

Best Local Similarity 51.4%; Pred. No. 2.1e-41;

Matches 93; Conservative 36; Mismatches 44; Indels 8; Gaps 2;

QY 1 MLTMDIIRGDHPTLRQAAELELPITKEEKTLLAMREFLVNSQDEETAKRYGLRSGVG 60

DB 13 LITMDIIRGNTLRQAAELELPITKEEKTLLAMREFLVNSQDEETAKRYGLRSGVG 72

QY 61 LAAPQINISKRMIAVLIPD--DGSCK-----SYDMLVNPNTVSHSVQEAYLPTGEGCL 112

DB 73 LAAPQIDVSKRIITAVLPNLPDKEGPNPKPEAYSWQEVLPNPKVSHSVQDAALSDGEGCL 132

QY 113 SVDDNVAGLVHRHNNKTIKAKDIEGNDIQLRLKGYPAIVFQHEIDHLNGVMFYDHDKDH 172

SQ Sequence 203 AA;

Query Match 45.1%; Score 446.5; DB 2; Length 203;
 Best Local Similarity 51.1%; Pred. No. 1.1e-40;
 Matches 92; Conservative 35; Mismatches 46; Indels 7; Gaps 2;

QY 1 MLTMKDIIRDGHTLRQKAAELEPLTKKEKETLIAMREFLVNSQDEETAKRYGLRSGVG 60
 DB 13 LIDMDIIRREGPYTLREVANDVTPLSDDEIILGKMLQFLHNSQDPVMAEKWGLRGGVG 72
 QY 61 LAAPQINISKRMIATLIP-----DGSGSYDY--MLVNPKIYSHSVQEAAYLPTGEGCLS 113
 DB 73 LAAPQLDISKRIIVAVENIVEGETPQPAYDELAIMYNPKIVSHSVQDAALGEGEGCLS 132
 QY 114 VDDNVAGLVHRHNKTIKAKDIEGNDIQLRLKGYPAIVFQHEIDHLNGVMFYDHIKDP 173
 DB 133 VDRNPGYVVRHARVTVDYFDKDGKHKRIKLGKYNISVQHEIDHNGIMFYDRINEKDP 192

RESULT 15

ABB53869
 ID ABB53869 standard; protein; 211 AA.

AC ABB53869;

XX 29-AUG-2003 (revised)

DT 16-MAY-2002 (first entry)

XX Lactococcus lactis protein def.

XX Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.

XX Lactococcus lactis; IL1403.

XX FR2807446-A1.

XX 12-OCT-2001.

XX 11-APR-2000; 2000FR-00004630.

XX 11-APR-2000; 2000FR-00004630.

XX (INRG) INRA INST NAT RECH AGRONOMIQUE.

XX Bolotine A, Sorokine A, Renault P, Ehrlich SD;

XX WPI; 2002-043418/06.

XX New nucleotide sequence useful in the identification or Lactococcus
 XX lactis and related species.

XX Claim 6; SEQ ID NO 571; 2504pp; French.

XX The present invention is related to a Lactococcus lactis nucleotide
 CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The nucleic
 CC acid sequence is useful in the detection and/or amplification of nucleic
 CC acid sequence, particularly to identify Lactococcus lactis or related
 CC species. The proteins of the invention are useful for the biosynthesis or
 CC biodegradation of a composition of interest. The invention helps research
 CC in lactic bacteria, particularly useful in the production of yogurt and
 CC cheese. Note: The sequence data for this patent is based on equivalent
 CC patent WO200177334 (published 18-OCT-2001) which is available in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences. (Updated on 29-AUG-2003 to
 CC standardise OS field)

SQ Sequence 211 AA;

Query Match 43.0%; Score 426; DB 5; Length 211;
 Best Local Similarity 49.2%; Pred. No. 2.2e-38;
 Matches 91; Conservative 36; Mismatches 46; Indels 12; Gaps 3;

QY 1 MLTMKDIIRDGHTLRQKAAELEPLTKKEKETLIAMREFLVNSQDEETAKRYGLRSGVG 60

DB 16 MISMDIIRREGPYTLREVANDVTPLSDDEIILGKMLQFLHNSQDPVMAEKWGLRGGVG 75
 QY 61 LAAPQINISKRMIATLIP-----DGSGSYDY--MLVNPKIYSHSVQEAAYLPTG 108
 DB 76 LAANQLGLLKKVIAVLIIPNEPEVDEGNEIPPKAYKREIMYNKAVVSHSVQDAAVEGG 135
 QY 109 EGCLSVDDNVAGLVHRHNKTIKAKDIEGNDIQLRLKGYPAIVFQHEIDHLNGVMFYDHI 168
 DB 136 EGCLSVDRVPGYVVRHARVTVEYVYKGEKKIKLKDFFALCVQHEIDHNGVMFYDHI 195
 QY 169 DKDHP 173
 DB 196 NNNDP 200

Search completed: March 31, 2004, 16:35:03
 Job time : 60 secs